ETS GRE®

Assess Ability. Predict Performance.

ETS GRE® Board Research Report
ETS GRE®–17-01
ETS RR–17-03

Perceptions and Uses of *GRE*® Scores After the Launch of the *GRE*® revised General Test in August 2011

David M. Klieger

Vinetha Belur

Lauren J. Kotloff

December 2017

The report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

****

****

As part of its educational and social mission and in fulfilling the organization's non-profit Charter and Bylaws, ETS has and continues to learn from and also to lead research that furthers educational and measurement research to advance quality and equity in education and assessment for all users of the organization's products and services.

GRE-ETS

PO Box 6000

Princeton, NJ 08541-6000

USA

To obtain more information about GRE programs and services, use one of the following:
Phone: 1-866-473-4373
(U.S., U.S. Territories*, and Canada)
1-609-771-7670
(all other locations)
Web site: www.gre.org

*America Samoa, Guam, Puerto Rico, and US Virgin Islands

GRE Board Research Report Series and ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

Perceptions and Uses of *GRE*® Scores After the Launch of the *GRE*® revised General Test in August 2011

David M. Klieger, Vinetha Belur, & Lauren J. Kotloff

Educational Testing Service, Princeton, NJ

This survey study investigated how graduate school admissions committees perceive and use the *GRE*® General Test and *GRE*® Subject Tests after the launch of the *GRE*® revised General Test in August 2011. These perceptions and uses impact the validity of the tests. Prior research about the perceptions and uses of the General Test and Subject Tests was last conducted in 2002 and, prior to that, in 2000 (for writing) and 1984. Therefore, even without test revisions, perceptions and uses of the GRE tests may have changed. Overall, by graduate discipline area, by graduate degree level sought, and by future career track preparation (research vs. professional), we examined online survey responses from 163 individuals involved in graduate school admissions. We did not find major changes in the use or perceived utility of the revised General Test or Subject Tests for admissions or funding decisions. General Test scores are valued in relation to other admissions information. Sometimes, valuation of these scores leads to practices proscribed by published GRE Program guidelines, such as the use of cut scores. Scores on the Subject Tests continue to be valued less than other admissions information. Perceptions and uses of the GRE tests often vary when decision makers consider applications from international applicants and underrepresented racial or ethnic minorities. That admissions committees can receive and compare scores reported on the former General Test score scales to scores reported on the revised General Test score scales has resulted in various practices, including the common use of percentile ranks to compare scores in ways discouraged by GRE Program guidelines. To the extent that decision makers use concordance information recommended by Educational Testing Service (ETS) to compare scores, they do so with mixed success.

Keywords GRE; revised; Subject Test; graduate school; admissions; decisions; decision making; score scale

doi:10.1002/ets2.12130

In August 2011, the *GRE*® revised General Test of Verbal Reasoning, Quantitative Reasoning, and Analytical Writing Skills (revised GRE) was launched. The revisions to the former version of the test were made to "[m]ore closely align with the skills needed to succeed in graduate and business school; [p]rovide more simplicity in distinguishing performance differences between candidates; [p]rovide more test taker–friendly features for an enhanced test experience; [and e]nhance test security" (Briel & Michel, 2014, p. 1.1.1). The changes to the verbal reasoning subtest included more emphasis on complex reasoning in lieu of decontextualized vocabulary knowledge. Antonyms and analogies were eliminated, and highlighting of relevant sentences was permitted on the computer-based version. The changes to the quantitative reasoning subtest also included more emphasis on reasoning, with greater focus on real-life situations and interpretation of data, the addition of an onscreen calculator, and questions that allow for numerical responses rather than choosing among preexisting options. The analytical writing subtest now elicits more specific answers to questions to reduce the use of canned responses by examinees. In addition, the test design changed from a computer-adaptive test (CAT) to a multistage adaptive model (MST) that is a compromise between a linear model and the CAT model. Furthermore, the score scales for the verbal and quantitative sections changed from 200 to 800 (in 10-point increments) to 130 to 170 (in 1-point increments) to make salient to score users the changes in test content and format as well as to minimize score users' possible misperception that small incremental differences between possible scaled scores are statistically or practically meaningful.

With considerations of validity and fairness issues in mind, Educational Testing Service (ETS, 2015b) has issued the *GRE Guide to the Use of Scores 2015–2016* that advises score users how to use and not use GRE scores. These guidelines, approved by the GRE Board, consisting of administrators and faculty with expertise in graduate and professional education, include recommendations to use multiple criteria (i.e., not just GRE scores) to make decisions, to refrain from the establishment and implementation of GRE cut scores, to refrain from the use of percentile ranks to compare test scores when the percentile ranks being compared are based on different populations, and to transition from the old score scale

Corresponding author: D. M. Klieger, E-mail: dklieger@ets.org

(200 – 800 in 10-point increments) to the new score scale (130 – 170 in 1-point increments) using ETS-provided score concordance information. Some of these guidelines were applicable prior to the launch of the GRE revised General Test (e.g., recommendations to use multiple criteria and to avoid cut scores; see ETS, 2010), but others are particular to the move to the revised GRE (e.g., to transition from the old to the new score scale). Furthermore, the last substantial investigations of perceptions and uses of GRE scores were undertaken 14 or more years ago (Powers & Fowles, 2000; Walpole, Burton, Kanyi, & Jackenthal, 2002, for writing), and in some ways they were not as comprehensive as the investigation prior to them 32 years ago (Oltman & Hartnett, 1984). Therefore, due to known changes to the GRE's content, format, and scoring in August 2011, as well as changes in perceptions and uses of GRE test scores that might have occurred without those changes, it is important to reconsider issues concerning the GRE's validity, defined as an evaluation of the proposed interpretations and uses of test scores (Kane, 2006). Understanding stakeholders' current perceptions and uses of GRE scores, including issues that admissions committees face as a result of the introduction of a new GRE test in August 2011, is integral to this reconsideration.

Overview of Past Research Regarding Perceptions and Uses of GRE General Test Scores

Perceptions and uses of GRE scores presumably have been linked to the goals of graduate and professional programs. These objectives have included maximizing achievement and productivity (e.g., the level of knowledge and skill level attained by students in the program, the amount of knowledge that students and the program disseminate through publishing and teaching undergraduates), enhancing program diversity (e.g., by race, gender, socioeconomic status), and promoting a congenial working environment (see Enright & Gitomer, 1989; Kuncel, Hezlett, & Ones, 2001; Kyllonen, 2008; Oltman & Hartnett, 1984; Walpole et al., 2002). Prior research on the former version of the GRE indicated that, in trying to achieve admission and funding goals, scores have not been the primary basis by which many graduate admissions committees make acceptance decisions (Oltman & Hartnett, 1984). Rather, scores often have been utilized as an additional consideration for applicants with weaker credentials or as one part of an intuitive selection process in which committees holistically examine all submitted materials to develop an overall impression of an applicant (Monahan, 1991; Powers & Fowles, 2000; Rem, Oren, & Childrey, 1987; Walpole et al., 2002).

Alternative uses for scores from past versions of the GRE test have also been documented. First, in a survey by Oltman and Hartnett (1984), half of the institutions requiring the GRE used the scores to make decisions regarding fellowship and assistantship awards (though they did not specify by what procedures). Second, there is evidence that, despite clear warnings against the practice in the guidelines regarding score use, admissions committees have used score cutoffs to filter out applicants (Walpole et al., 2002). Only 10% of institutions surveyed by Oltman and Hartnett indicated admission committee adherence to the guidelines then in existence. Third, there has been variation in how scores from the GRE quantitative and verbal sections are emphasized by department; science, technology, engineering, and mathematics (STEM) departments place greater emphasis on quantitative scores, but social sciences, humanities, and education departments place greater emphasis on verbal scores (Walpole et al., 2002). These practices have been substantiated in ETS validity research (Schneider & Briel, 1990). Last, there are those programs that have chosen to not require or recommend submission of GRE scores at all (Oltman & Hartnett, 1984).

Potential Challenges in Score Interpretation

Furthermore, it is imperative that admissions committees that receive revised GRE scores are able to interpret and utilize the information provided by those scores in an informed manner. Although the majority of committee members are faculty who assess their students, they nevertheless are not necessarily experts in assessment for admissions and funding. Research has consistently documented the phenomenon of lay misinterpretation of statistical information (Carpenter & Shah, 1998). Furthermore, there is evidence to support the claim that even experts have difficulty interpreting statistical information (Ross, 1990). Chi (2006) reported that experts often have difficulty making accurate predictions on the performance of novices in their fields, (e.g., predicting how well a student will perform in graduate school). These potential challenges may arise for interpretation and use of GRE scores, particularly given that admissions committees can receive GRE scores on the new score scale (from the revised version of the GRE), GRE scores on the former score scale (from the former version of the GRE), or both—and both from the same applicant with multiple sets of GRE scores as well as different applicants who have taken different versions of the test. Given that (a) GRE scores are reportable for up to 5 years

after their respective test dates and (b) the former version of the *GRE*® General Test ceased operation in August 2011, the additional challenge of comparing scores on different scales began to disappear in August 2016. Meanwhile, and as previously mentioned, the guidelines (ETS, 2015b) expressly encourage the use of concordance information to transition from the old to the new score scale.

Tradeoffs in Achieving Admissions and Funding Objectives

Decision makers may perceive tradeoffs among the goals they have in admissions and funding. These tradeoffs may have implications for how the GRE is perceived and used. For example, an admissions committee might reject an application from someone with great promise for technical competence (someone with high GRE scores and a high undergraduate grade point average) due to financial limitations or because a letter of recommendation indicates that the applicant is unable to work well with others. Increasingly scarce federal and state funding for higher education and limits on the tuition for which students are able and willing to pay may increase the frequency and impact of some of these tradeoffs.

Some graduate and professional programs have effectuated over the past few decades policies that consider the gender and racial or ethnic diversity of their student bodies. According to Bowen and Bok (1998), several reasons may account for this. First, in order to identify and promote individuals of high potential, institutions have taken into account groups that have historically experienced prejudice and discrimination and continue to experience fewer opportunities for economic advancement (cf. Fix & Struyk, 1993; Riach & Rich, 2002). Second, the diversity of the student body exposes students to a wide variety of perspectives and experiences, a learning process that arguably is essential as our society itself is becoming increasingly diverse. Finally, there is the idea that the members of the most educated and influential professions of an increasingly diverse society should reflect that diversity, lest the country return to previously experienced lower levels of inclusion (Bowen & Bok, 1998; Katzenbach & Marshall, 1998). These considerations also may also cause admissions committees to weight the GRE (and other admissions information) differently for groups that score lower on it.

In addition, the use in admissions of a standardized cognitive assessment such as the GRE may make the achievement of certain types of diversity more challenging. Simultaneously maximizing the accuracy of predicting success outcomes (e.g., graduate GPA) and the racial and ethnic diversity of those admitted (especially the inclusion of Black and Hispanic applicants) when cognitive assessments are used in selection has been a long-standing challenge (see Pyburn, Ployhart, & Kravitz, 2008). The GRE General Test and *GRE*® Subject Tests are predictive of academic performance (Klieger, Cline, Holtzman, Minsky, & Lorenz, 2014; Kuncel & Hezlett, 2007; Kuncel et al., 2001; Liu, Klieger, Bochenek, Holtzman, & Xu, 2016). However, for U.S. citizens, the average standardized test scores of White and Asian examinees have generally exceeded the average scores of Black and Hispanic examinees by at least one half to a full standard deviation (see, e.g., ETS, 2014; Gallagher, Bridgeman, & Cahalan, 2000). To the extent that admissions systems utilize cut scores (especially top-down selection), score differences of this magnitude can have a substantial impact on the percentage of Hispanic and African American examinees admitted (see Sackett & Wilk, 1994). Despite the usually greater accuracy of using statistical methods in predicting academic achievement (Kuncel, Klieger, Connelly, & Ones, 2013; Kuncel, Klieger, & Ones, 2014), diversity considerations motivate some admissions committees to use holistic methods to facilitate diversity goals (Foderado, 2009) or to weight the GRE (and perhaps other admissions information) differently for groups that score lower on it. Presumably, recent trends in U.S. federal law that limit the availability of affirmative action (*Fisher v. University of Texas at Austin*, 2013, 2015; *Gratz v. Bollinger*, 2003; Grutter v. Bollinger, 2003; *Schuette v. Coalition to Defend Affirmative Action*, 2014) encourage this approach.

The GRE Subject Tests

As the name implies, the GRE Subject Tests assess acquired knowledge and skills in specific subject areas. According to the GRE Subject Tests website (ETS, 2015a), these tests are intended to supplement other evidence of an individual's qualifications for graduate study and can identify the strengths and weakness of an individual's preparation in the intended course of study preparation. Subject Tests have been part of the GRE Program since the early 1950s (Conrad, Trismen, & Miller, 1977). By the late 1970s, Subject Tests were offered in 20 areas. Currently, Subject Tests are offered in the following seven areas:

- Literature in English
- Mathematics

- Biochemistry, cell and molecular biology
- Biology
- Chemistry
- Physics
- Psychology

Research has shown Subject Tests to be strong predictors of graduate school performance. In their meta-analysis of the validity of the GRE and undergraduate GPA as predictors of graduate school success, Kuncel et al. (2001) found the Subject Tests to have higher operational validity than the verbal, quantitative, or analytic subtests of the GRE. Historically, however, the Subject Tests have been given lower priority by graduate school admissions committees compared with the GRE General Test and other qualifications (Oltman & Hartnett, 1984; Walpole et al., 2002). Reasons given for lesser use of the GRE Subject Tests have included (a) they indicate what applicants' undergraduate major or minor programs taught rather than what an applicant can learn or do in graduate school; (b) too many desirable candidates lack knowledge and skills measured by the Subject Tests, often because their undergraduate majors and minors differed from the graduate field in question; (c) admissions committees do not appreciate the knowledge that Subject Tests require; and (d) the number of available test subject areas has decreased (Oltman & Hartnett, 1984; Walpole et al., 2002). Given the historical integration of the Subject Tests into the GRE Program and their relatively higher validity yet lessened role in graduate and professional school admissions, it is important to provide an update on stakeholders' perceptions and uses of the Subject Tests.

In summary, the current study investigated, from the perspective of those who make admissions and funding decisions for graduate students, the following issues overall—by discipline (STEM vs. social science vs. humanities), by degree program level (doctorate vs. master's), and by program career track preparation (professional track vs. research track):

1. the perceived utility of the GRE revised General Test for decision-making, including for admissions decisions and awarding fellowships;
2. the perceived utility of the GRE Subject Tests for decision-making;
3. how GRE scores are valued in relation to other admissions information;
4. how the need to achieve multiple goals at the same time, such as maximizing diversity and maximizing predictive validity, affect how committee members' perceive and use the GRE; and
5. how the coexistence of former and new score scales for the quantitative and verbal reasoning sections is affecting the perception and use of scores, including how the receipt of multiple sets of GRE General Test scores from the same or different score scales is handled.

Method

Survey Instrument

In order to determine how the GRE revised test is perceived and used, we administered an online survey to graduate and professional school faculty and administrators. We developed the survey by starting with the comprehensive questionnaire used by Oltman and Hartnett (1984) and then modifying it to incorporate considerable input from the GRE Program and graduate deans and provosts on the GRE Board as well as new questions specific to the transition from the former version of the GRE General Test to the current version. The resulting survey contained logic branching that sometimes routed respondents to subsequent questions based on their prior answers. A template of the survey appears in Appendix A. Of the 42 items in the survey, 15 items asked about the role the GRE test played in graduate admissions decisions (e.g., as a ranking criterion, in matching applicants with faculty, etc.) as well as admissions-related decisions (e.g., awards and funding, course placement), whereas six items asked about how admissions processes are impacted by ETS's introduction of the new scoring scale in August 2011. An additional two items asked about selectiveness of the participant's institution and the institution's policies on admitting students with disabilities. One item empirically tested participants' abilities in converting between the old and new score scales with the use of concordance tables that ETS made available on its website to graduate institutions and programs. Finally, 18 background items assessed participants' gender, age, work experience, citizenship, title, program, department, admissions committee service, and participant familiarity with guidelines established by the GRE Board for the use of GRE scores (e.g., ETS, 2010, 2015b).

Procedure

An e-mail blast containing a link to the online survey was sent to graduate school administrators and faculty who had participated in prior graduate school research projects with ETS Research and Development. The e-mail asked them for referrals to other administrators and faculty who they believed had participated in or were familiar with the graduate school admissions process. This first data collection occurred from January to May 2014 and garnered 90 sets of survey responses. After response and referral rates slowed dramatically from May to December 2014, e-mail recipients were encouraged to directly forward the invitation to participate in the survey to other administrators and faculty. The data collection then continued until March 2015 and resulted in another 73 sets of responses. Ultimately, there were 163 respondents to the survey. One hundred and fifty-seven participants who completed the survey were classifiable into one of three discipline areas of STEM, social sciences (SSc), or arts–humanities (AH). The breakdown by discipline area, our primary variable of interest in this study, was 41% ($n = 64$) STEM, 34% ($n = 54$) SSc, and 25% ($n = 39$) AH, which reflected the maximum sample sizes that we could achieve with a reasonable expenditure of resources. Regarding perceptions and uses of GRE test scores, we also were interested in examining whether differences exist between master's programs (25%, $n = 40$) versus doctoral programs (75%, $n = 121$), as well as programs that are geared toward professional (23%, $n = 37$) versus research (77%, $n = 126$) career tracks.

Analysis

In this study, both the overall sample size ($n = 163$, with even smaller *n*s for subsamples) and some sampling bias, resulting from the fact that a few small clusters of respondents came from the same school or program, dissuaded us from conducting statistical significance testing. Instead, the process we used to identify whether a result was practically significant and thus noteworthy was based on graphical analyses and then effect sizes. First, we (a) created graphical representations of the data, (b) visually inspected the graphs for the highest degrees of difference or nondifference across the variables of interest, and (c) considered whether such differences or nondifferences made both common and conceptual sense. A difference was believed to exist when one category visually stood out above the rest. Similarly, if no category stood out above the rest during visual inspection of the graphs, the result was described as nondifference, even if small percent differences existed between the categories.

Once this methodology was applied and a noteworthy result was believed to exist, the strength of the difference was analyzed using mean (M) differences and Cohen's d effect sizes for continuous variables and percent differences and relative risk ratios (RRRs) for categorical variables. A Cohen's d of 0 indicates no effect (i.e., no difference between groups' means) whereas a nonzero Cohen's d indicates an effect, or difference. The further away Cohen's d is from 0, the larger the effect. The sign (positive or negative) of the Cohen's d value indicates which group has the higher mean. The RRR is a ratio of two probabilities (one representing the group of interest and the other representing the comparison group); an RRR of 1 indicates no difference between the two groups. The distance away from 1 is an indicator of the magnitude of the difference between the two groups. An RRR of less than or greater than 1 indicates group differences exist. An RRR of less than 1 indicates the group of interest is less likely than the comparison group to endorse a particular response, whereas an RRR of greater than 1 indicates the group of interest is more likely than the comparison group to endorse a particular response. For example, if 100% of Group A replies "yes," but only 25% of Group B does, then the RRR is 100% ÷ 25%, or a 4 times greater likelihood that Group A will reply "yes" than will Group B.

We did not rely on Cohen's (1988) guidelines to assess the practical significance of effect size values. As Cohen himself suggested, his guidelines risk providing false certainties. In the end, we concluded that our overall analytic approach was optimally reasonable. We believe that not only will readers inevitably make inferences and generalizations based on anything reported, but they need to do so in order to make decisions related to the revised GRE. We also believe that this report strikes a reasonable balance between psychometric orthodoxy and practical need.

Results

Participant Characteristics

The majority of the participants in the sample were White (74%), U.S. citizens (94%), and almost equally represented in gender (46% female, 49% male, 5% not reported). A small number (6%) of participants identified as being of Latino

Table 1 Participant Characteristics

Variable	Level	*n* (%)
Gender	Female	75 (46%)
	Male	80 (49%)
	Prefer not to answer	8 (5%)
Latino/Hispanic background	Yes	10 (6%)
	No	143 (88%)
	Prefer not to answer	10 (6%)
Race	White/Caucasian	120 (74%)
	Asian	12 (7%)
	Black/African American	6 (4%)
	Biracial/multiracial	4 (2%)
	Do not identify with these	7 (4%)
	Prefer not to answer	14 (9%)
Citizenship	United States only	141 (89%)
	United States dual	8 (5%)
	Non-United States	10 (6%)
Worked outside academia	Yes	50 (31%)
	No	109 (67%)
	Prefer not to answer	4 (2%)
Admissions service	Committee chair	90 (55%)
	Committee member	119 (73%)
	Provided advice	18 (11%)
	Received advice	6 (4%)
	Other	7 (4%)
Program type[a]	Professional master's	22 (13%)
	Research master's	18 (11%)
	Professional doctorate	15 (9%)
	Research doctorate	106 (65%)
	Other	2 (1%)
Admissions policies[a]	Not competitive	0 (0%)
	Somewhat competitive	13 (8%)
	Moderately competitive	61 (37%)
	Very competitive	59 (36%)
	Extremely competitive	30 (18%)
Relevant GRE Subject Tests available[a]	Yes	72 (44%)
	No	85 (52%)
	Missing/unknown	6 (4%)

[a]Refers to the program in which the participant was most involved in making admissions decisions.

or Hispanic background. Most (73%) had served as an admissions committee member. Approximately half (55%) had served as committee chairs, 11% had provided advice to an admissions committee, and 4% had received advice from such a committee. With regard to the type of program that participants were most involved in while serving on an admissions committee, nearly two thirds (65%) indicated they were involved in research doctorate programs, followed by professional master's (13%), research master's (11%), professional doctorate (9%), and other (1%). Thirty-seven percent of participants said admissions policies in those programs are moderately competitive, 36% said very competitive, 18% said extremely competitive, 8% said somewhat competitive, and 0% said not competitive. See Table 1 for more about sample characteristics.

Institution Characteristics

The 163 participants in the final analysis came from 31 different U.S. higher education institutions from the South (35%), East (19%), West (19%), and Midwest (26%), with the plurality from the South. Based on the Carnegie classification system, the majority of the institutions were public institutions (77%), large (74%), residential campuses (71%), and with high or very high research activities (77%) taking place. Most institutions (65%) offered comprehensive doctoral graduate instruction. See Table 2 for additional institutional characteristics.

Table 2 Institutional Characteristics

Variable	Level	*n* (%)
Region	South	11 (35%)
	East	6 (19%)
	West	6 (19%)
	Midwest	8 (26%)
Institution type	Public	24 (77%)
	Private	7 (23%)
Size	Large	23 (74%)
	Medium	7 (23%)
	Small	1 (3%)
Campus type	Residential	22 (71%)
	Nonresidential	9 (29%)
Focus	Doctoral/research	1 (3%)
	Master's colleges and universities	6 (19%)
	Research university (high research activity)	11 (35%)
	Research university (very high research activity)	13 (42%)
Graduate instruction	Comprehensive doctoral, no medical/veterinary	12 (39%)
	Comprehensive doctoral with medical/veterinary	8 (26%)
	Doctoral, humanities/social sciences dominant	1 (3%)
	Doctoral, professional dominant	4 (13%)
	Doctoral, STEM dominant	3 (10%)
	Postbaccalaureate professional	1 (3%)
	Postbaccalaureate with arts and sciences	1 (3%)
	Single doctoral (education)	1 (3%)

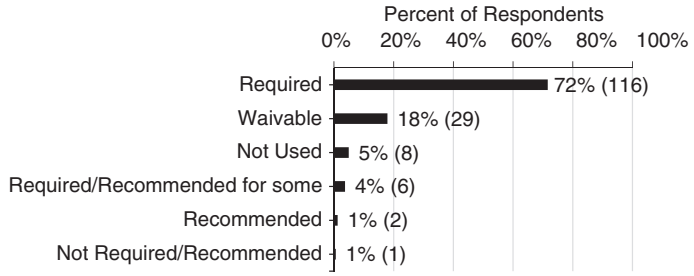

Figure 1 University requirement policy for the GRE General Test.

Survey Main Results

The results from the survey can be divided into two main phenomena of interest: (a) ways the revised GRE is used in admissions decisions and admissions-related decisions and (b) issues faced by admissions committees as a result of the introduction of a new GRE test in August 2011.

Use of the GRE in Admissions and Funding Decisions

Participants were surveyed on whether their institution requires the revised GRE test; the importance of revised GRE test scores compared to other credentials; and whether and how revised GRE scores are used for decisions such as matching students to faculty, course placement, and assistantships or fellowships.

Are the GRE General Test and GRE Subject Tests Required?

Figures 1 – 8 display requirement policies related to GRE General Test and Subject Tests overall and by discipline, degree program level, and career track. Note that in these and future figures, frequencies (*n* values) appear in parentheses adjacent to the percentages that represent them.

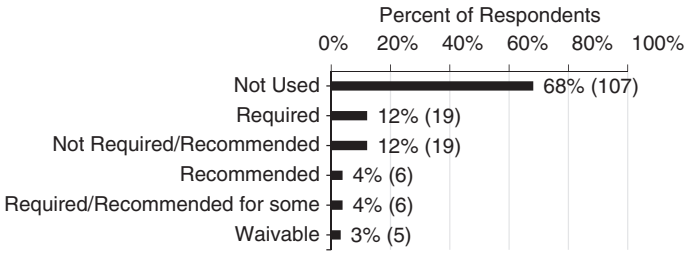

Figure 2 University requirement policy for the GRE Subject Tests.

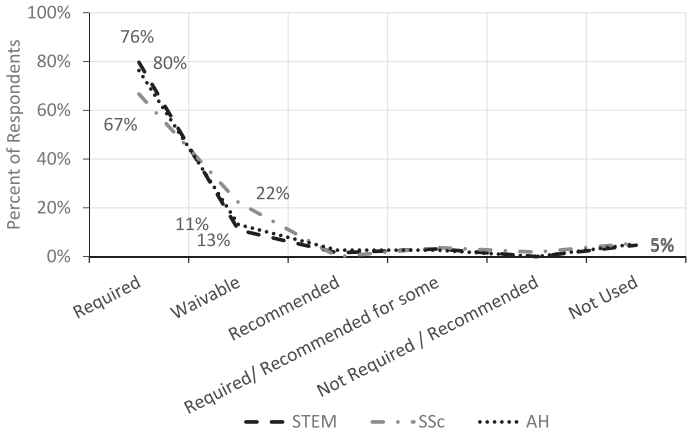

Figure 3 University requirement policy for the GRE General Test, by discipline.

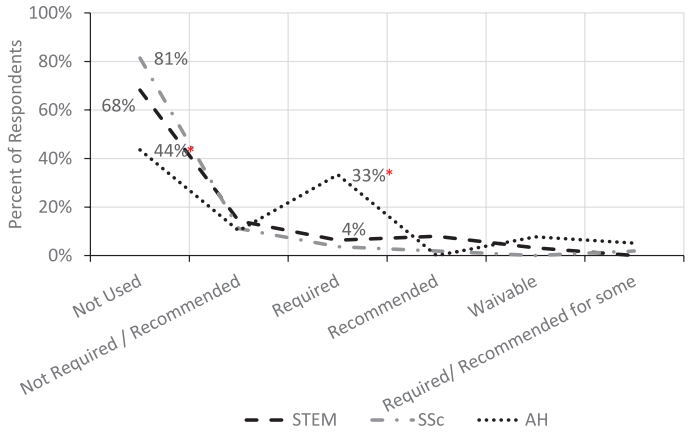

Figure 4 University requirement policy for the GRE Subject Tests, by discipline The percentages for the use of GRE Subject Tests for AH programs possibly are inflated because of an overrepresentation of English literature programs in our sample (18 out of 39, or 46%, of AH programs) for which one of the AH Subject Tests (literature in English) is available (see the red qualifying asterisks).

Although the GRE revised General Test is widely used, the GRE Subject Tests are not. As shown in Figure 1, 72% of respondents indicated the GRE revised General Test is required for all candidates, 18% said it is required but could be waived in some circumstances, 4% said it is required or recommended for some applicants, 1% said it is recommended for all applicants, fewer than 1% said it is neither required nor recommended but would be considered if submitted, and 5% said it is not used at all. In contrast, Figure 2 shows that 68% of respondents indicated GRE Subject Tests are not used in admissions decisions, 12% said they are not required or recommended but would be considered if submitted, 12% said they are required for all candidates, 4% said they are required or recommended for some candidates, 4% said they are recommended for all candidates, and 3% said they are required but could be waived in some circumstances.

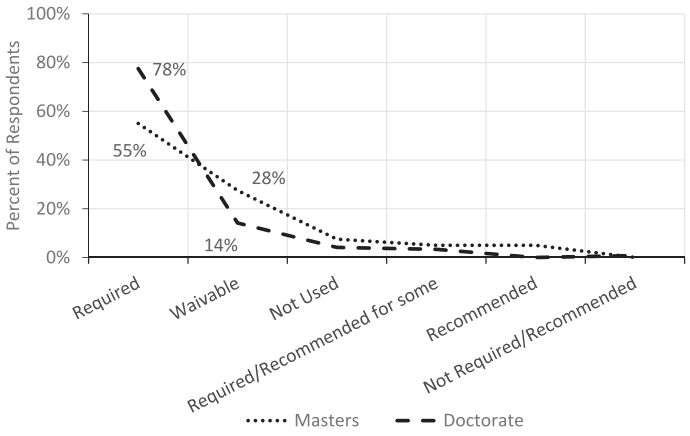

Figure 5 University requirement policy for the GRE General Test, by degree level.

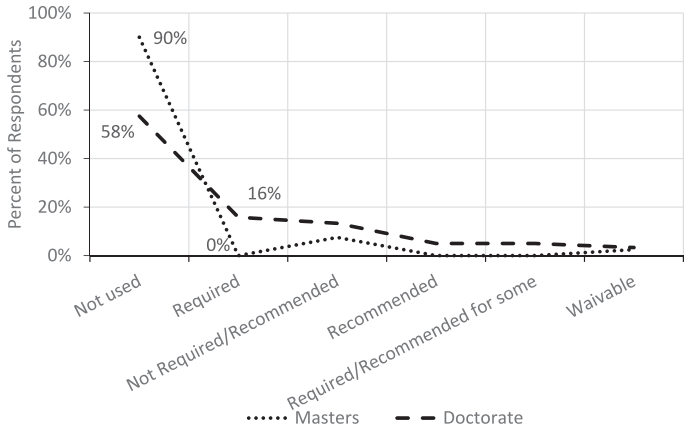

Figure 6 University requirement policy for the GRE Subject Tests, by degree level.

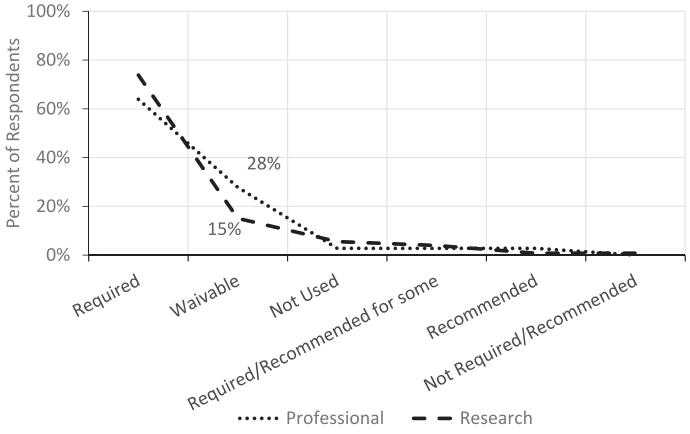

Figure 7 University requirement policy for the GRE General Test, by career track.

With respect to differences by discipline (STEM, SSc, AH), Figure 3 shows that STEM and AH respondents were more likely than SSc respondents to report that the GRE General Test is required at their institution (80% and 76% vs. 67%, RRR = 1.2 and 1.1, respectively) and were less likely than SSc respondents to say the test could be waived (11% and 13% vs. 22%, RRR = 0.5 and 0.6, respectively). There were no differences across the disciplines in whether the GRE General Test is (a) recommended for all applicants, (b) required or recommended for some degree applicants, or (c) not

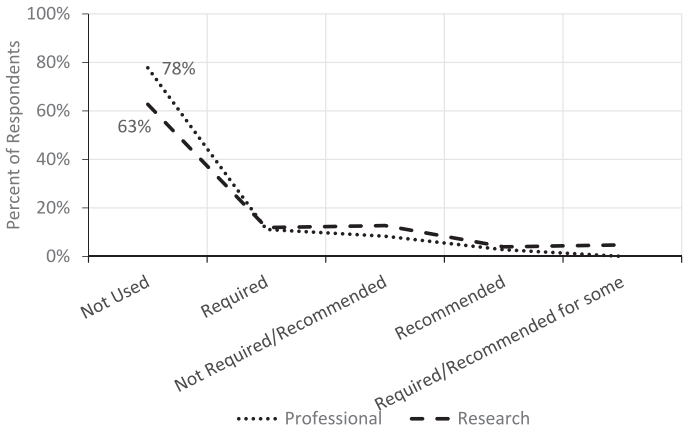

Figure 8 University requirement policy for the GRE Subject Tests, by career track.

required or recommended but considered if submitted. However, as Figure 4 conveys, SSc and STEM respondents were less likely than AH respondents to report that GRE Subject Tests scores are required for all applicants as part of their program's policy (4% and 6% vs. 33%; RRR = 0.1 and 0.2, respectively). Similarly, SSc and STEM respondents were more likely than AH respondents to report that GRE Subject Tests are not used in admissions policy decisions (81% and 68% vs. 44%, RRR = 1.9 and 1.6, respectively). Given that five out of seven operational Subject Tests are in STEM disciplines (mathematics; biochemistry, cell and molecular biology; biology; chemistry; and physics), a greater use of Subject Tests in STEM fields was expected. However, the numbers for the use of GRE Subject Tests for AH programs possibly are inflated because of an overrepresentation of English literature programs in our sample (18 out of 39, or 46%, of AH programs) for which one of the AH Subject Tests (literature in English) is available; hence, the red qualifying asterisks in Figure 4. Only 10% of respondents from non-English literature AH programs indicated that their programs require the GRE Subject Tests; only 19% of them neither require nor recommend the Subject Tests (yet will consider them if scores are submitted). Given that none of the remaining response categories received any responses, we cannot determine the percentage of non-English literature AH programs that simply do not use the GRE Subject Tests. Given the lack of AH Subject Tests other than literature in English, we surmise that the percentage of non-English AH programs that do not use the GRE Subject Tests is very high, and higher than for STEM programs (with five STEM-related Subject Tests).

With respect to differences by degree program level (master's or doctorate), master's programs generally rely less on revised GRE scores than do doctorate programs. Regarding the GRE General Test, participants affiliated with master's programs were less likely than participants affiliated with doctoral programs to say the GRE General Test is required for all applicants (55% to 78%, RRR = 0.7) and more likely than participants affiliated with doctoral programs to say that, although generally required, it can be waived in certain circumstances (28% to 14%, RRR = 1.9). See Figure 5. Participants from master's programs were also less likely to say the GRE Subject Tests are required for all applicants (0% vs. 16%, RRR = 0) and were more likely to say the GRE Subject Tests are not used in admissions decisions (90% vs. 58%, RRR = 1.6). See Figure 6.

With respect to differences by career track (professional or research), participants from professional tracks were more likely to report that the GRE revised General Test can be waived sometimes (28 vs. 15%, RRR = 1.8). See Figure 7. Participants from professional-track programs were more likely to report not using the GRE Subject Tests (78% vs. 63%, RRR = 1.2). See Figure 8.

The response rate for the survey question about the importance of listed reasons for a program not requiring or recommending GRE test scores for admissions (Question 31 in Appendix A) was so low ($n \leq 8$) that we do not report findings for this question.

How Are GRE Scores Used Along With Other Credentials?

Figures 9 – 12 show how admissions committees may use revised GRE test scores relative to other credentials. As Figure 9 shows, overall most respondents (78%) reported that revised GRE scores are used as one factor in a holistic review of applicants' files. That practice is consistent with ETS's GRE guidelines (ETS, 2015b), which advise that GRE score users

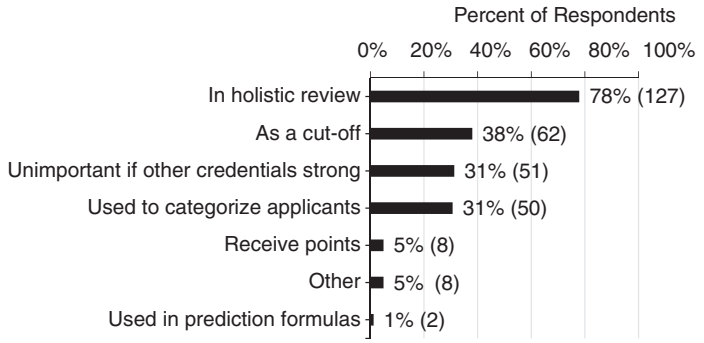

Figure 9 How GRE Scores are used relative to other credentials.

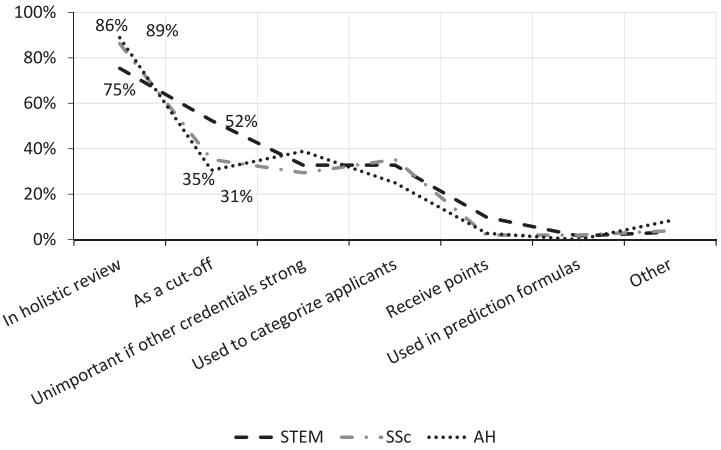

Figure 10 How GRE Scores are used relative to other credentials, by discipline.

utilize multiple information sources to make admissions and student funding decisions because of fairness concerns and the multidimensionality of the determinants of graduate and professional student success. Thirty-eight percent of respondents reported that applicants with a GRE score below a particular cutoff are not considered. This approach is inconsistent with the guidelines, which expressly warn against the use of cut scores. Other reported methods of GRE score use did not conflict with the guidelines. Thirty-one percent of respondents reported that when other credentials are strong, GRE test scores are not important, but for applicants with weaker credentials, test scores are expected to compensate; and 31% reported that GRE scores are used to categorize applicants as being *probable*, *possible*, or *unlikely* before other credentials are reviewed. Few to no respondents (5%) indicated they assign points to each applicant's GRE test scores and other credentials based on how important they are perceived to be. Few to no respondents (1%) indicated they use prediction formulas (e.g., regression equations) based on GRE test scores and other credentials of previous applicant groups. The eight respondents who indicated they utilize other ways of using GRE scores in admissions reported the following: Two participants indicated their institution used no cutoff score, and therefore even participants with the lowest GRE scores could be accepted or that individual programs could waive the cutoff if desired; two participants reported that very high scores could be used to secure funding outside of the department and at the college level; one participant reported that analytical writing and quantitative reasoning scores were used to identify the areas where students needed assistance; and three participants repeated what was already endorsed in other categories or were off topic.

With respect to differences by discipline, STEM respondents, relative to AH and SSc respondents, were more likely to report that applicants with a GRE score below a particular cutoff are not considered (52% vs. 31% and 35%, RRR = 1.7 and 1.5, respectively). Again, the use of cut scores is inconsistent with guidelines advising about how GRE scores should and should not be used (ETS, 2015b). STEM respondents were less likely than AH and SSc respondents to report that scores are one factor in a holistic review (75% vs. 89% and 86%, RRR = .8 and .9, respectively). See Figure 10.

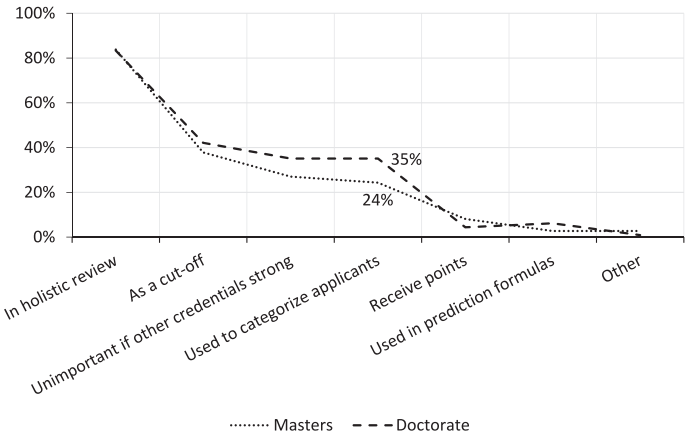

Figure 11 How GRE scores are used relative to other credentials, by degree level.

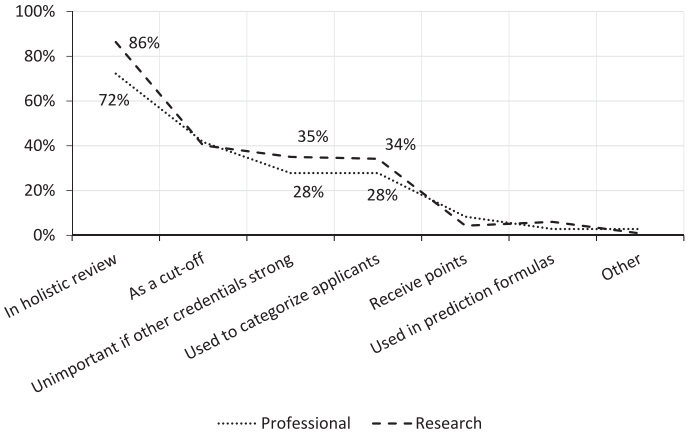

Figure 12 How GRE scores are used relative to other credentials, by career track.

With respect to differences by degree program level, doctorate program respondents are more likely than master's program respondents to use GRE scores to categorize applicants first before other credentials are used (35% vs. 24%, RRR = 1.5). See Figure 11.

With respect to differences by career track, respondents from research programs were more likely than respondents from professional programs to report GRE scores are (a) used as one factor in a holistic review, (b) unimportant if other credentials are strong, and (c) used to categorize applicants first before other credentials are used (86% vs. 72%, 35% vs. 28%, 34% vs. 28%, RRR = 1.2, 1.3, and 1.2 respectively). See Figure 12.

How Important Are GRE Test Scores Compared to Other Credentials?

Figures 13 through 16 display results related to the level of importance placed on GRE revised General Test scores and 25 other criteria during the admissions decision-making process overall and across discipline, degree level, and program type. A score of 0 (*low*) indicates the criterion is not used in admissions decisions; a score of 1 indicates the criterion is not very important in admissions decisions; a score of 2 indicates the criterion is moderately important; a score of 3 indicates the criterion is very important; and a score of 4 (*high*) indicates the criterion is extremely important for admissions decisions. Scores were averaged across respondents. Arrows indicate the ranked position of GRE Verbal Reasoning, GRE Quantitative Reasoning, and GRE Analytical Writing scores.

Figure 13 illustrates the overall relative importance of GRE tests in admissions. In general, GRE Verbal Reasoning test scores, GRE Quantitative Reasoning test scores, and GRE Analytical Writing test scores were perceived to be of moderate importance in admissions decisions relative to other admissions information and had the 12th ($M = 2.4$, SD = .98), 14th

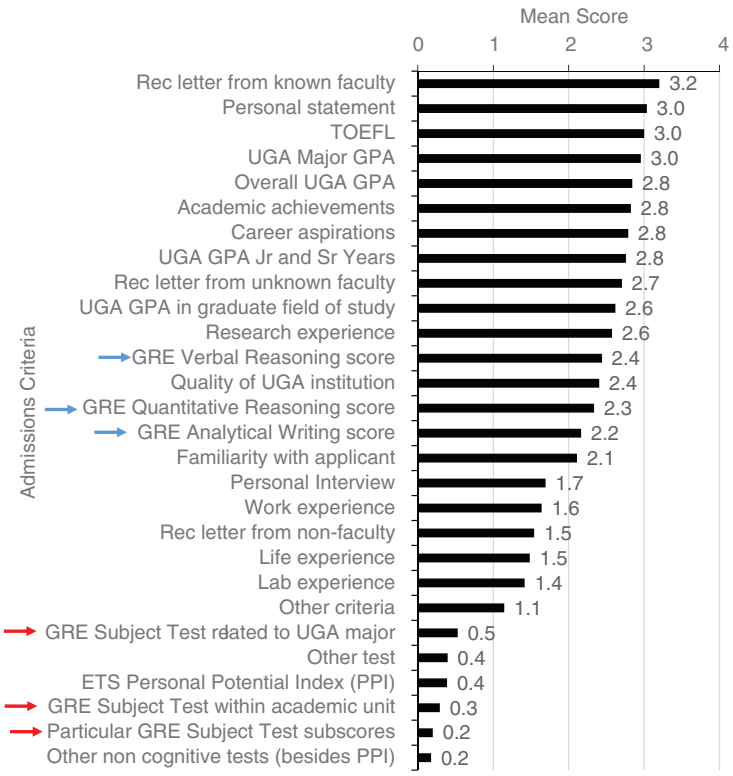

Figure 13 Decreasing importance of admissions criteria. The blue arrows indicate where the three sections of the GRE General Test fall in the rank order. The red arrows indicate where the GRE Subject Tests fall in the rank order.

($M = 2.3$, SD = 1.24), and 15th ($M = 2.2$, SD = 1.09) highest means (out of 28 criteria assessed), respectively. GRE Subject Tests particular to the academic unit ($M = 0.3$, SD = .75), particular GRE Subject Tests subscores ($M = 0.2$, SD = .56), and other noncognitive tests ($M = 0.2$, SD = .51), except for the *ETS*® Personal Potential Index (*ETS*® PPI), were the three least important criteria for admissions decisions. Of the 44 respondents who reported other criteria were used in admissions, the following criteria were reported by 28 respondents (16 participants did not indicate what other criterion is used): writing samples ($n = 13$), research potential or research authorship ($n = 3$), evidence of time management ($n = 3$), leadership experiences ($n = 2$), study abroad experiences ($n = 2$), background check or legal–criminal history ($n = 2$), creative accomplishments ($n = 1$), communication skills ($n = 1$), and diversity ($n = 1$).

With respect to differences by discipline, Figure 14 (second black bold arrow) illustrates how STEM respondents are likely to give more importance to lab experience during admissions decisions relative to AH and SSc respondents ($M = 2.3$ vs. 0.1 and 1.2; $d = 2$ and 0.8, respectively). STEM respondents and SSc respondents are likely to give more importance to GRE Quantitative Reasoning scores during admissions decisions relative to AH respondents (first black bold arrow; $M = 2.9$ and 2.6 vs. 1.0; $d = 2.0$ and 1.6, respectively). STEM and SSc respondents are likely to give less importance to other criteria during admissions decisions relative to AH respondents (third black bold arrow; $M = 0.6$ and 0.9 vs. 2.3, $d = -1.3$ and −0.9, respectively). The other criteria deemed to be important by AH respondents were mostly writing samples (65%) or study abroad experience (10%).

As expected, Figure 15 indicates, based on an almost uniformly higher trend line for doctorate programs, that these programs generally are more selective with the admissions criteria than are master's programs. With respect to differences by degree level, relative to participants affiliated with doctorate programs, participants affiliated with master's program admissions committees give less importance to research experience ($M = 1.8$ vs. 2.8, $d = -1.1$), academic achievements ($M = 2.3$ vs. 3.0, $d = -1.3$), a personal interview ($M = 1.2$ vs. 1.9, $d = -.7$), lab experience ($M = 1.0$ vs. 1.6, $d = -.4$), the GRE Subject Tests related to undergraduate major ($M = 0.1$ vs. 0.7, $d = -0.6$), and other criteria ($M = 0.5$ vs. 1.3, $d = -.9$).

With respect to career track (professional vs. research), participants affiliated with professional tracks are likely to give more importance to overall undergraduate GPA ($M = 3.1$ vs. 2.8, $d = 0.5$) and recommendation letters from nonfaculty

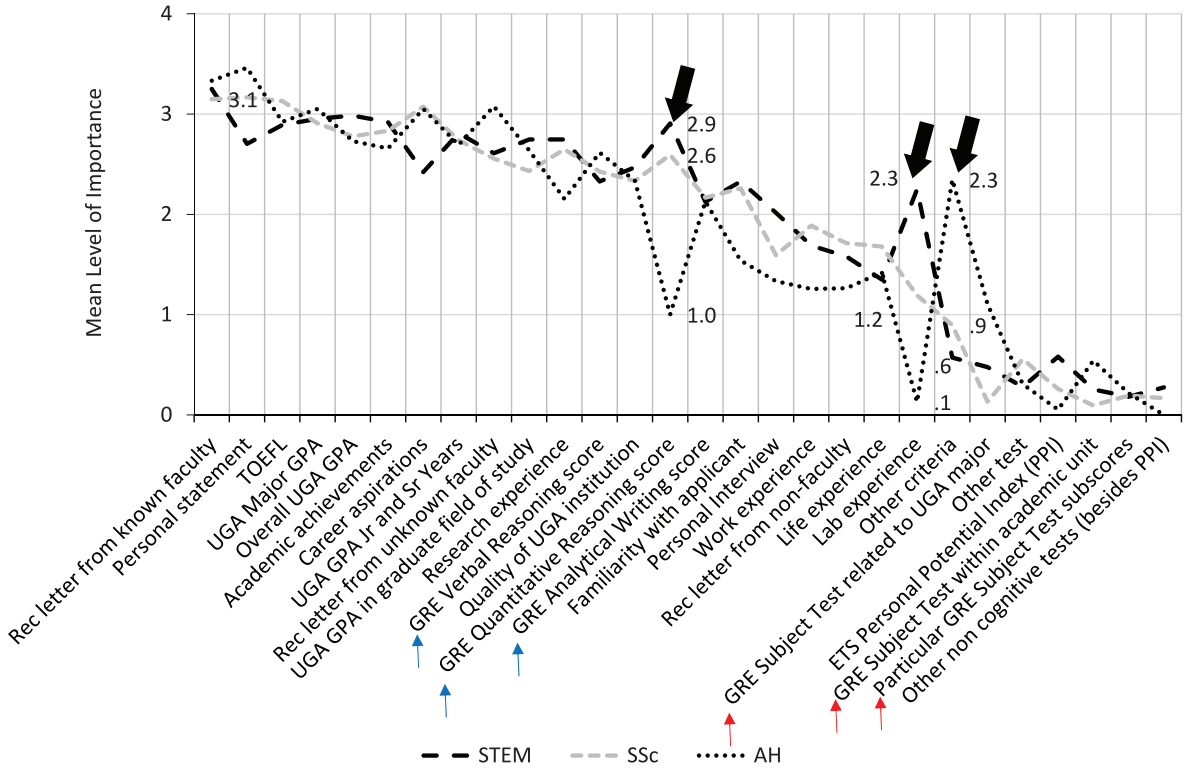

Figure 14 Decreasing importance of admissions criteria by discipline. The three bold black arrows signify criteria with the greatest differences by discipline. The blue and red arrows indicate where the three sections of the GRE General Test and the Subject Tests, respectively, are located within the various admissions criteria.

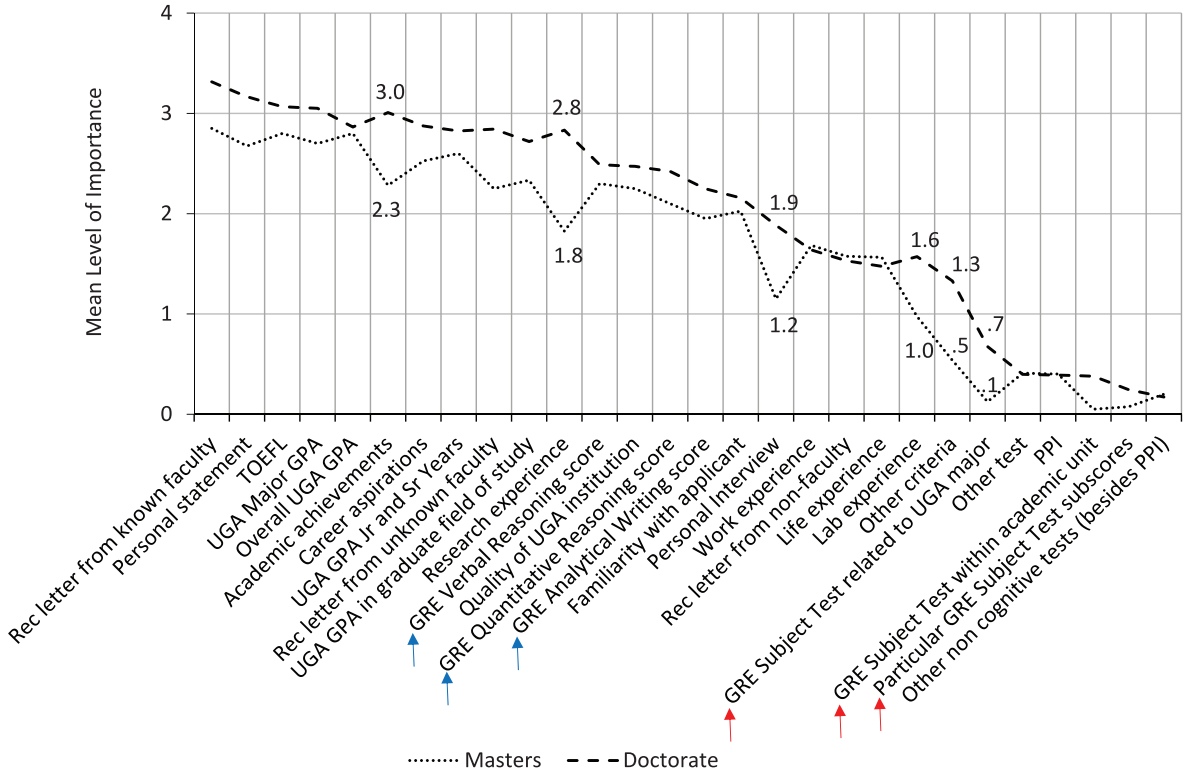

Figure 15 Decreasing importance of admissions criteria by degree level. The blue and red arrows indicate where the three sections of the GRE General Test and the Subject Tests, respectively, are located within the various admissions criteria.

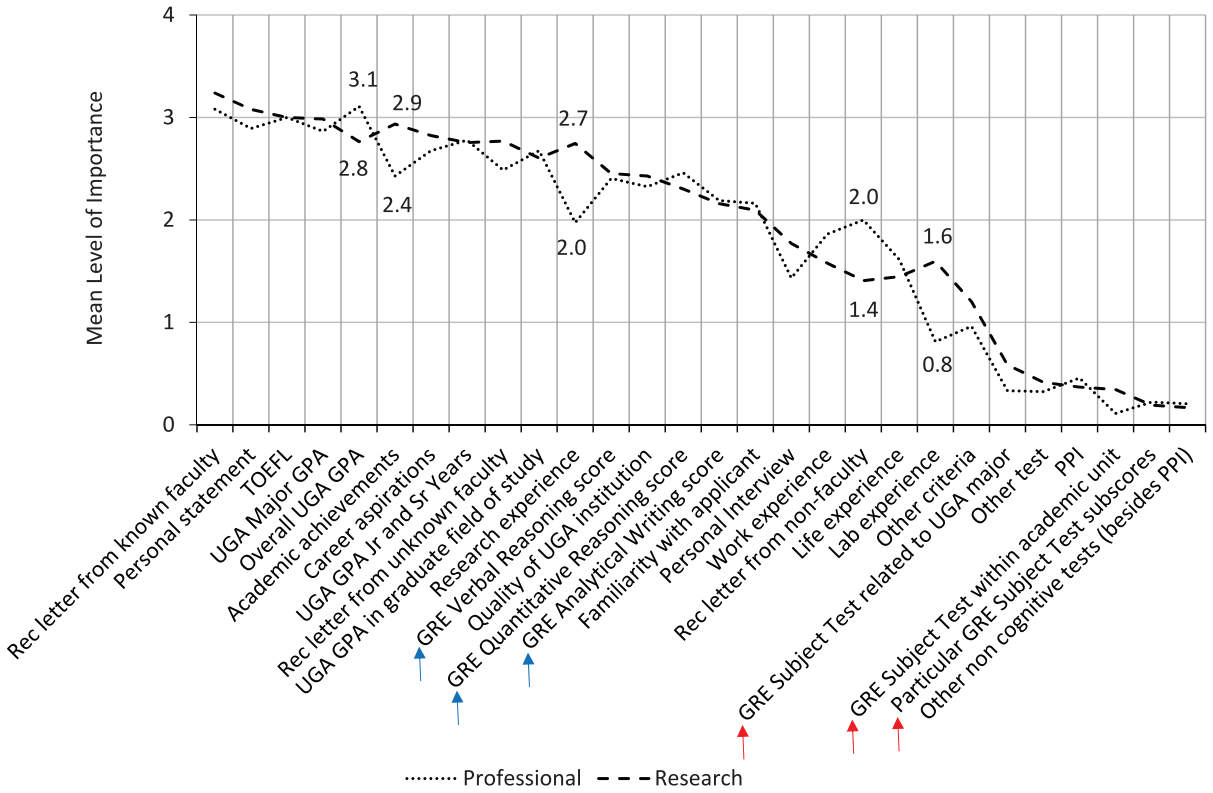

Figure 16 Decreasing importance of admissions criteria by career track. The blue and red arrows indicate where the three sections of the GRE General Test and the Subject Tests, respectively, are located within the various admissions criteria.

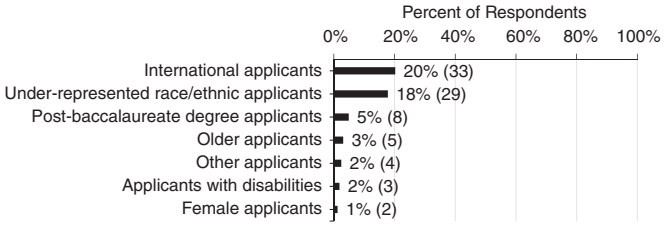

Figure 17 Change in relative importance of admissions criteria based on candidate background.

($M = 2.0$ vs. 1.4, $d = 0.7$) than participants affiliated with research tracks. See Figure 16. Participants affiliated with professional tracks are less likely to give importance to academic achievements ($M = 2.4$ vs. 2.9, $d = -0.6$), research experience ($M = 2.0$ vs. 2.7, $d = -0.7$), and lab experience (0.8 vs. 1.6, $d = -0.6$) than participants affiliated with research tracks. See Figure 16.

Does the Relative Importance of Admissions Criteria Change Based on Subgroup?

Figures 17 through 20 indicate whether the relative importance of the 28 criteria would change if the applicant was of a particular background such as an international, underrepresented ethnic minority, older, or female subgroup.

As Figure 17 shows, overall 20% of participants indicated that the relative importance of the 28 criteria would change if the applicants are international students, and 18% said it would change for underrepresented racial or ethnic groups. International students may have different prior educational experiences; the standards by which they have been previously judged via grades and recommendation letters may differ; some programs might wish to separately consider their English speaking and listening skills in addition to their reading comprehension and writing skills. Previously mentioned diversity goals might explain variance in how information is weighed for racial and ethnic minority candidates. Few participants

said it would change for applicants who had a postbaccalaureate degree (5%), for older applicants (3%), for those with disabilities (2%), for females (1%), and for other subgroups (2%). Of the four respondents who indicated the relative importance would change for "other subgroups," these were the subgroups reported: men ($n = 1$), nondegree-seeking ($n = 1$), working professionals ($n = 1$), and applicants without undergraduate degrees in the graduate study area ($n = 1$).

With respect to discipline, STEM respondents relative to AH and SSc respondents (14% vs. 28% and 24%, RRR = 0.5 and 0.6, respectively) were less likely to report that the relative importance of the 28 criteria is different for international applicants (see Figure 18).

Figure 19 illustrates differences in relative importance by degree level. Doctoral program participants felt that the importance given to the above criteria changed for international students and underrepresented minorities but not for (in descending order of whether importance of criteria would change) older applicants, postbaccalaureate applicants, female applicants, applicants with disabilities, and other groups. Master's program participants reported a similarly ordered trend, but they were less likely than doctoral participants to say the relative importance of the criteria would change for internationals (30% vs. 43%, RRR = 0.8) and underrepresented ethnic minorities (17% vs. 39%, RRR = 0.5) and more likely to say criteria would change for older applicants (13% vs. 3%, RRR = 4.5) and other applicants (13% vs. 2%, RRR = 9.1).

With respect to career track, respondents from research-track programs were more likely to report the relative importance of criteria would change for underrepresented race or ethnic groups (37% vs. 27%, RRR = 1.4) and less likely to report criteria would change for other applicants (2% vs. 12%, RRR = 0.2). See Figure 20.

Uses of GRE General Test Scores

Figures 21 – 28 display results related to the various admissions-related and other uses of the GRE revised General Test, such as in selecting applicants who can help to achieve certain goals of the program or institution, or in determining funding or placement decisions. Few respondents ($n < 10$) indicated they use the GRE Subject Tests for any kind of funding, advising, or placement purposes, so those results have been omitted. For Figures 21 through 24, a score of 0 (*low*) indicates the GRE revised General Test is not applicable toward achieving the goals; a score of 1 indicates the GRE revised General Test is not very important in achieving the goal; a score of 2 indicates the GRE revised General Test is moderately important; a score of 3 indicates very important, and a score of 4 (*high*) indicates the GRE revised General Test is extremely important in achieving the goal.

Overall, the GRE revised General Test is most important for the goal of selecting applicants who would be able to handle graduate school coursework ($M = 2.6$, $SD = 1.09$). Figure 21 shows a distinct drop in the level of importance of the GRE revised General Test for achieving all other goals. Of the 14 participants who indicated they use the GRE revised General Test to meet some other goal than what was listed, the following other goals were reported: grad school or other administration requires it as a way of rating the program institutionally and nationally ($n = 6$); obtaining university fellowships ($n = 3$); selecting applicants who will impact the real world ($n = 1$); as a first step, broad-stroke picture of the candidate ($n = 1$); as an indication of overall intelligence and academic skill ($n = 1$); and as a way of determining if coursework can be handled ($n = 2$).

As shown by Figure 22, 44% of respondents indicated they use GRE revised General Test scores for awarding assistantships and fellowships. Few respondents indicated they use the GRE revised General Test for academic advising (6%), placement of students in courses (2%), as a comprehensive exam or other graduation requirement (2%), or for another purpose (6%). Less than 5% of all participants indicated they use the GRE Subject Tests for awards, placement, academic advising, as a comprehensive exam or other graduation requirement, or for another purpose.

The almost uniformly higher trend lines for STEM in Figures 23 and 24 indicate that STEM programs more strongly perceived GRE General Test scores to be useful to a broad spectrum of objectives in comparison to SSc and AH programs. With respect to differences by discipline, SSc and especially STEM respondents, relative to AH respondents, felt the role of GRE General to be more important in selecting students who will be able to learn important skills outside the classroom, such as the ability to run special software or lab equipment that facilitates research ($M = 1.6$ and 1.3 vs. 0.6, $d = 1$ and 0.7, respectively). See Figure 23. STEM respondents, relative to SSc and AH respondents, were more likely to use the GRE General Test scores for making decisions about assistantship and fellowship awards (56% vs. 35% and 43%, RRR = 1.6 and 1.3, respectively). They were also more likely than AH respondents to report using the GRE revised General Test for academic advising (11% vs. 0%). See Figure 24.

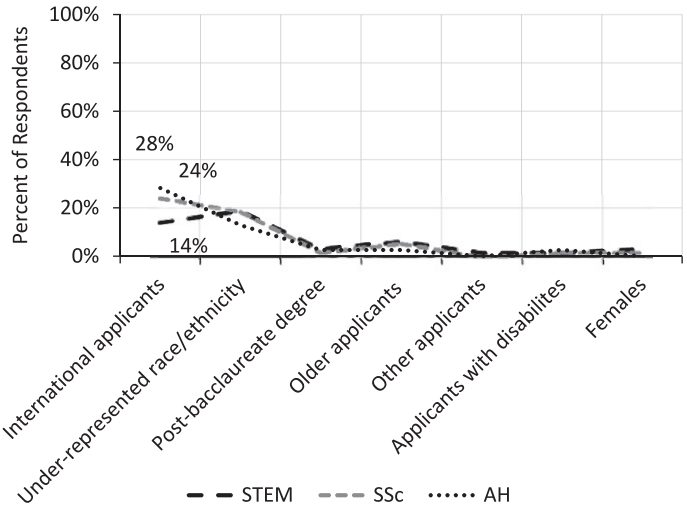

Figure 18 Change in relative importance of admissions criteria based on candidate background and discipline.

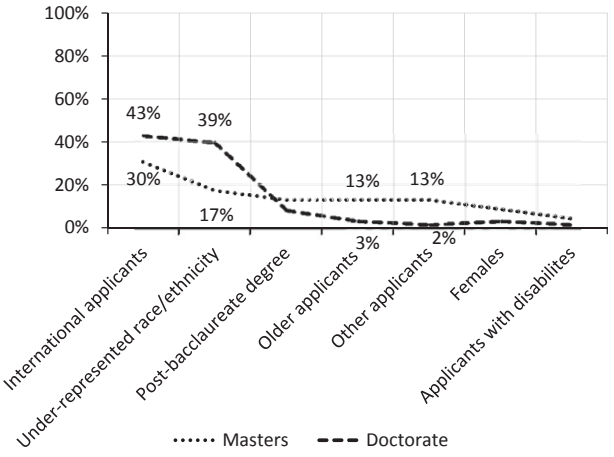

Figure 19 Change in relative importance of admissions criteria based on candidate background and degree level.

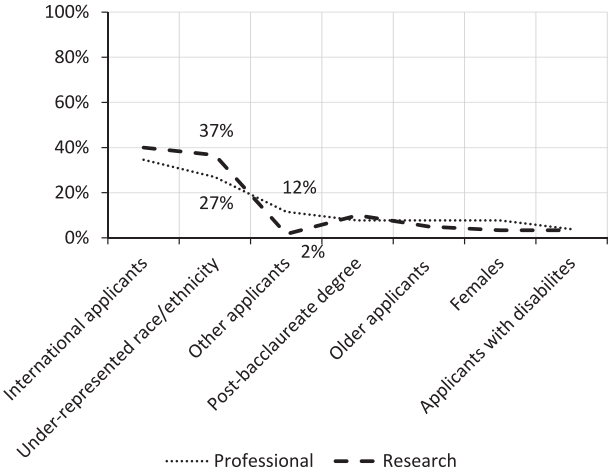

Figure 20 Change in relative importance of criteria, by subgroup and career track.

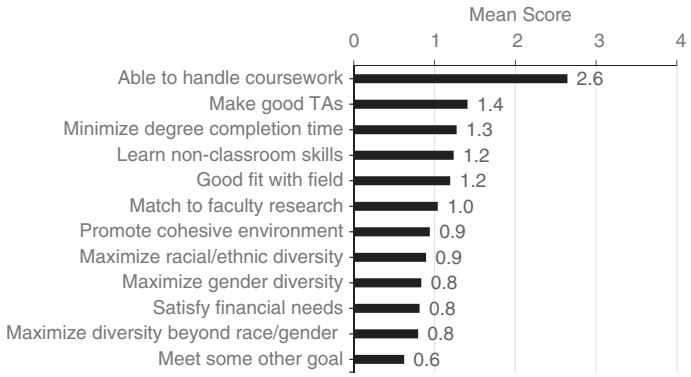

Figure 21 Perceived utility of the GRE General Test in selecting candidates who meet institutional or program goals.

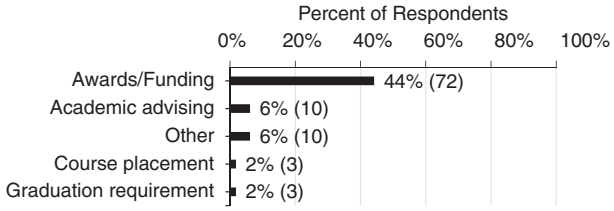

Figure 22 Use of the GRE General Test for funding, advising, and placement decisions.

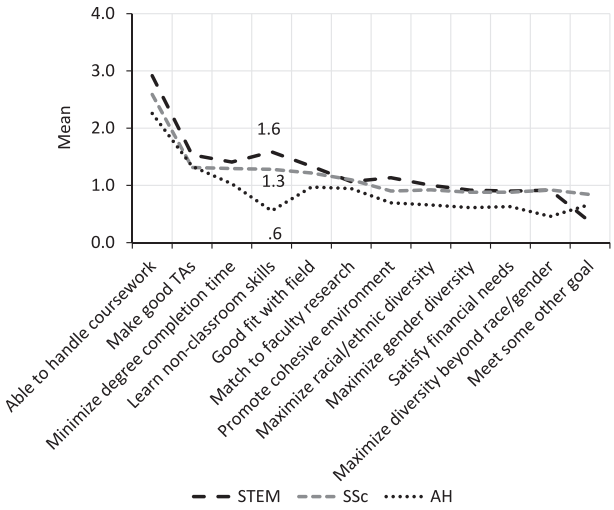

Figure 23 Perceived utility of the GRE General Test in selecting candidates who meet institutional or program goals, by discipline.

With respect to differences by degree level, participants in master's programs felt the GRE General Test is less important in achieving certain goals than participants in doctoral programs; hence the overall lower trend line for master's program respondents (see Figure 25). In particular, they felt the GRE revised General Test is less important in selecting good TAs (1.0 vs. 1.5, $d = -0.5$) and matching applicants' research interests with faculty expertise (0.6 vs. 1.2, $d = -0.5$) (see Figure 25). Master's participants are less likely than doctoral participants to use the GRE revised General Test for awarding assistantships and fellowships (35% vs. 52%, RRR = 0.7), but both groups are equally as unlikely to use the GRE revised General Test for purposes of academic advising, placing students in courses, as a comprehensive exam, or for another purpose (see Figure 26). The fact that the master's-level programs felt that GRE utility is lower for achieving these goals was unsurprising, but unexpectedly, 35% of the master's-level respondents saw the GRE General Test as useful for awards decisions, and master's-level respondents in general perceived it as being at least nominally relevant for a goal such as minimizing time to degree completion.

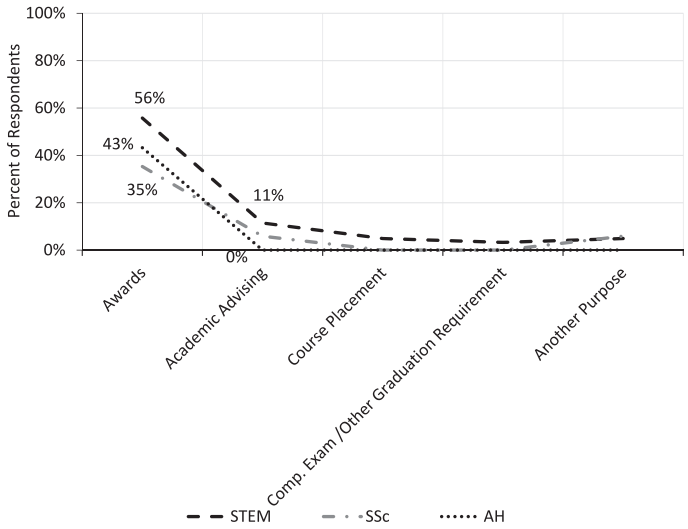

Figure 24 Use of the GRE General Test for funding, advising, and placement decisions, by discipline.

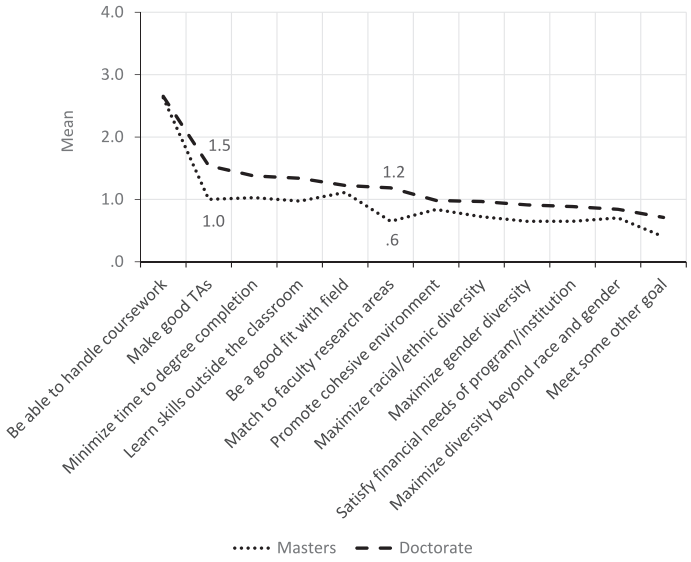

Figure 25 Perceived utility of the GRE General Test in selecting candidates who meet institutional or program goals, by degree level.

With respect to differences by career track, both research and professional programs perceived the GRE revised General Test as being most useful to identify those who are able to handle coursework (see Figure 27). Other potential purposes were seen as far less relevant. Participants from professional-track programs were less likely than participants from research-track programs to use the GRE General Test for awarding assistantships and fellowships (25% vs. 54%, RRR = 0.5). See Figure 28.

Issues Arising From the New GRE Scale

Figures 29 through 44 display results related to new GRE score scale and resulting test scores, including whether and how the new scale has changed the way test scores are perceived and used in admissions (Figure 29), whether preference remains for the former score scale (Figures 30 through 32), how multiple sets of scores from a single applicant are handled by admissions committees (Figures 33 through 36), how test scores from both the former GRE and the revised GRE from a single applicant are handled by admissions committees (Figures 37 through 40), and whether participants can convert scores between two different scales. Additionally, before examining issues related to the introduction of the scoring scale

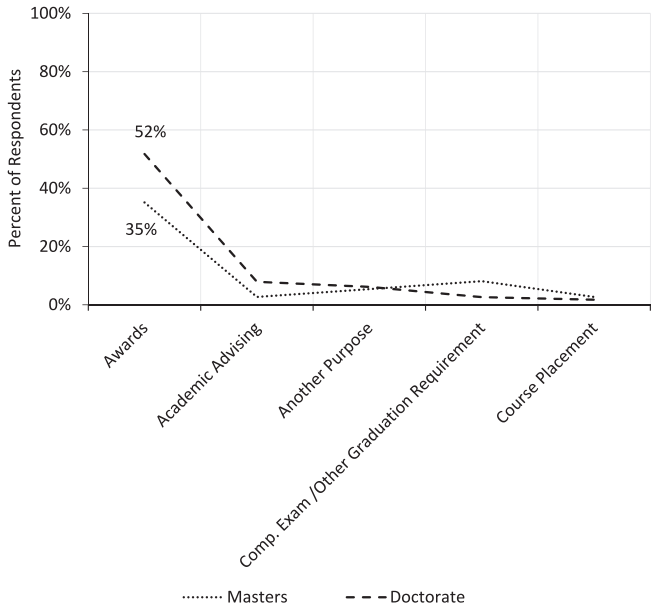

Figure 26 Use of the GRE General Test for funding, advising, and placement decisions, by degree level.

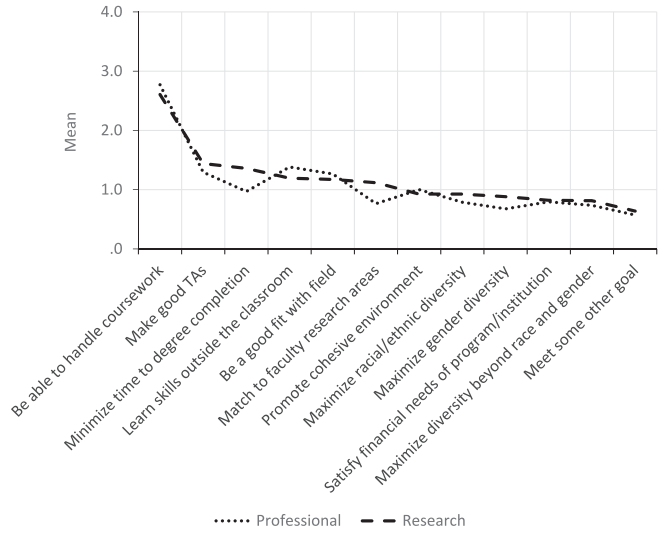

Figure 27 Perceived utility of the GRE General Test in selecting candidates who meet institutional/ program goals, by career track.

on the revised GRE, we first wanted to know how many participants were familiar with the guidelines from the GRE Board on the use of test scores.

Familiarity With GRE Board Guidelines on the Use of Test Scores

Most participants (67%) were unfamiliar with the guidelines. In terms of subgroup differences, SSc respondents were more likely to be familiar with the guidelines than AH respondents (41% vs. 16%, RRR = 2.5, where STEM = 33%), but no other notable subgroup differences were found by degree level or career track.

Has the New Scale Changed the Way Test Scores Are Used?

Figures 29 through 32 display results pertaining to whether and how the new scale has changed the way test scores are used in admissions.

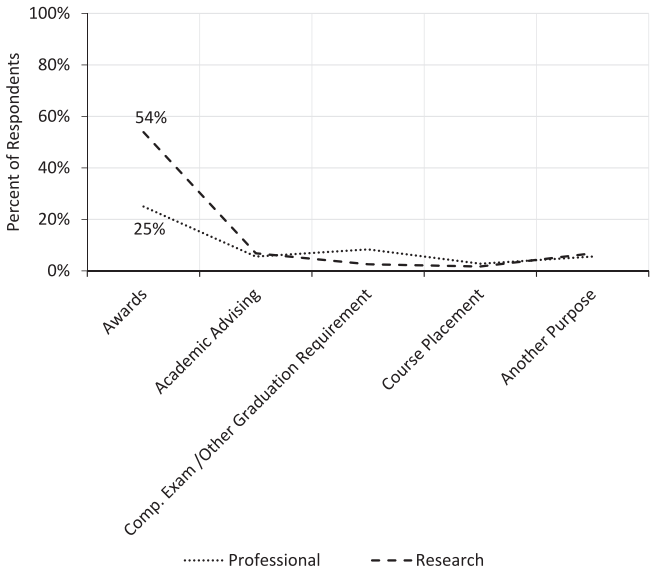

Figure 28 Use of the GRE General Test for funding, advising, and placement decisions, by career track.

Figure 29 describes how the new scale on the revised GRE affected use of GRE scores. Ninety-three percent of respondents reported that the revised scale has not changed the way in which they use GRE scores. A few respondents (3%) indicated that GRE scores play a bigger role in admissions as a result of the new scale; a few (1%) said that GRE scores play a smaller role in admissions; and 3% said that the new scale has affected their use in other ways. Thirty-one participants commented on the new scale. Fifteen (almost half) said they now use the percentile scores rather than the raw scores, whereas eight (about one quarter) did not answer the question directly but criticized the new scoring system. One said:

> I have found the new scoring system to be needlessly confusing and obfuscating. Since 2011, it has been FAR more difficult to easily evaluate scores. The matter is made only the more confusing by the way that older scores have been "converted" to the new system — without giving the corresponding percentile score, however, which is THE most useful metric. It caused quite a bit of confusion at the beginning and it makes it more difficult to compare those with the old version vs. the new.

Three other participants said they now convert to the old scale, and five were either unaware of the change, had no experience with it, or their comments were off topic.

With respect to differences by discipline, degree level, and career track, there were no practical differences in terms of how the new scale impacted the use of GRE scores.

Does Preference Exist for the Old Versus New Scoring Scale?

Figures 30 through 33 display results pertaining to whether there is a preference for the former or revised score scale.

In general, in connection with how receiving scores from the former and revised GRE for the same applicant or across different applicants influence their use of GRE scores, a majority of respondents (80%) indicated that they do not let the version of the test affect admissions or funding decisions, 13% said they always give preference to scores from the revised GRE, and 7% or fewer give preference to scores from the old version or give preference to one or the other version at certain times (see Figure 30).

With respect to differences by discipline, SSc respondents were less likely than AH respondents to report that preference is always given to the revised GRE (10% vs. 20%, RRR = 0.50). See Figure 31. With respect to degree level, there were no practical subgroup differences observed. With respect to career track, respondents from professional-track programs were less likely than those from research-track programs to report that the version of the test doesn't affect decisions (71% vs. 83%, RRR = .9), having slightly more preference for the revised test (see Figure 32).

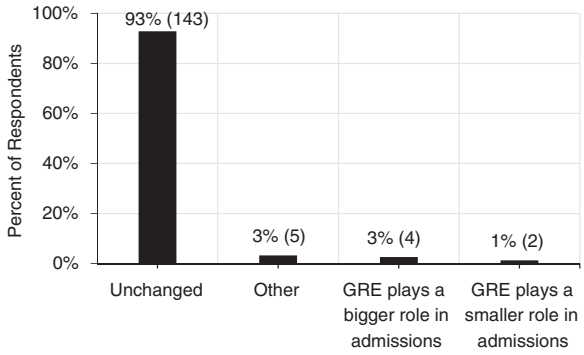

Figure 29 The new scale's impact on the way GRE scores are used.

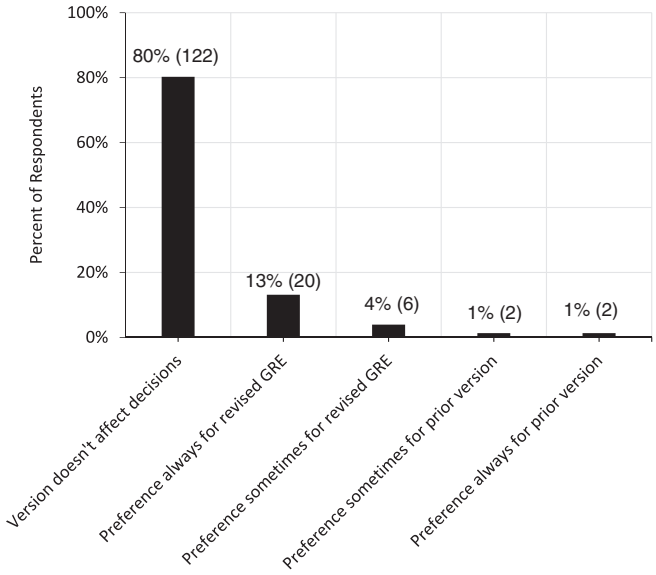

Figure 30 Preferences for old GRE scores versus revised GRE scores when receiving both scores from the same applicant or across different applicants.

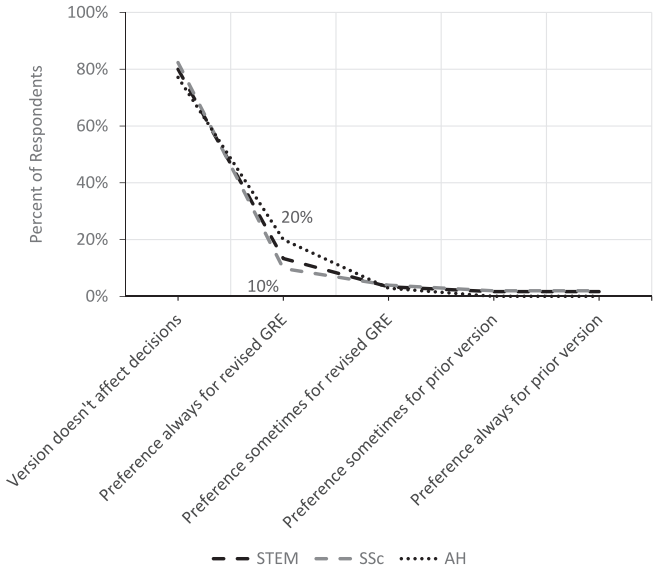

Figure 31 Preferences for old GRE scores versus revised GRE scores, by discipline.

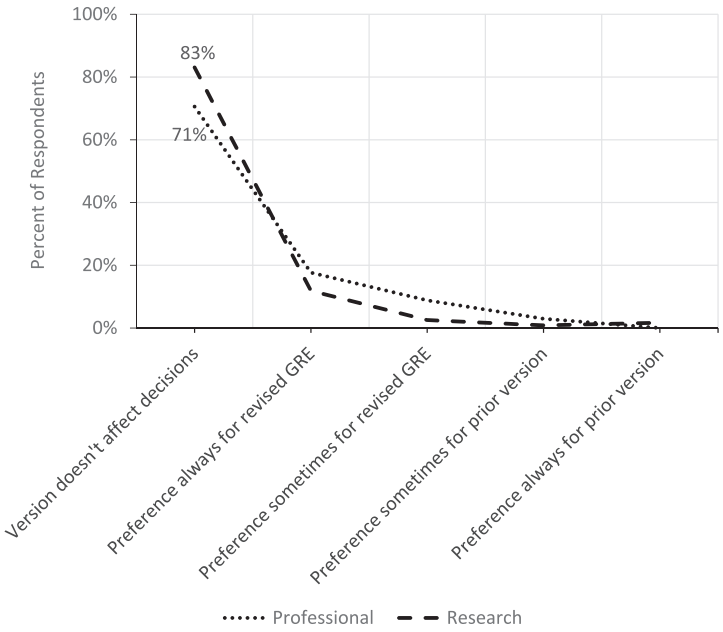

Figure 32 Preferences for old GRE scores versus revised GRE scores, by career track.

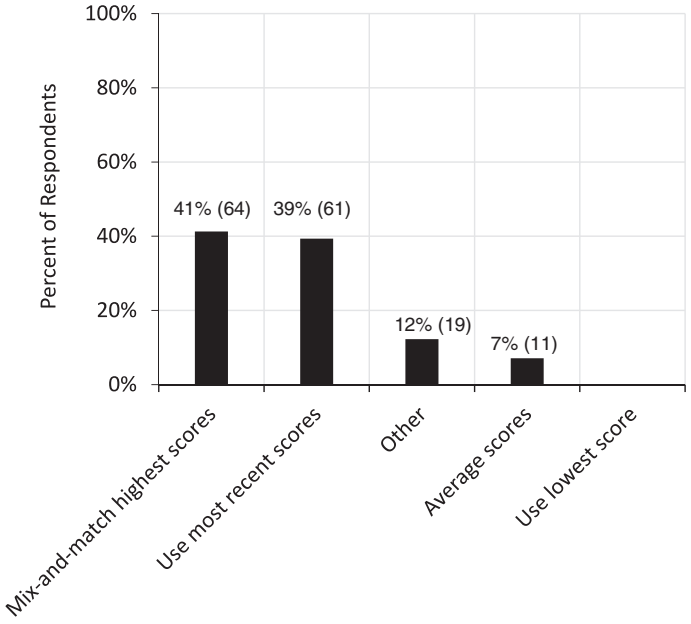

Figure 33 How admissions committees use multiple sets of GRE scores.

How Are Multiple Sets of Test Scores From the Same Applicant Handled?

Figures 33 through 36 display results pertaining to how admissions committees handle multiple sets of scores from a single applicant.

Forty-one percent of respondents reported that when their programs receive multiple sets of GRE scores from an applicant, their programs use the highest score from each GRE section, whereas 39% of respondents reported they use the scores from the most recent test administration (see Figure 33). A few respondents (7%) reported they average the scores from each GRE section. Twelve percent of respondents said multiple sets of GRE scores are handled in some other way, such as by looking at all of the scores but giving weight to most recent scores; using all of the scores as an estimate of the student's abilities; or by looking for an explanation of score discrepancies. No respondents reported using the lowest score

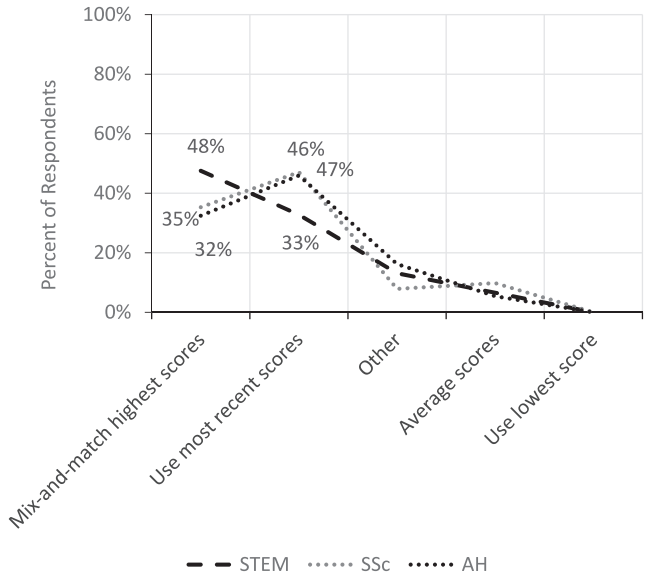

Figure 34 How admissions committees use multiple sets of GRE scores, by discipline.

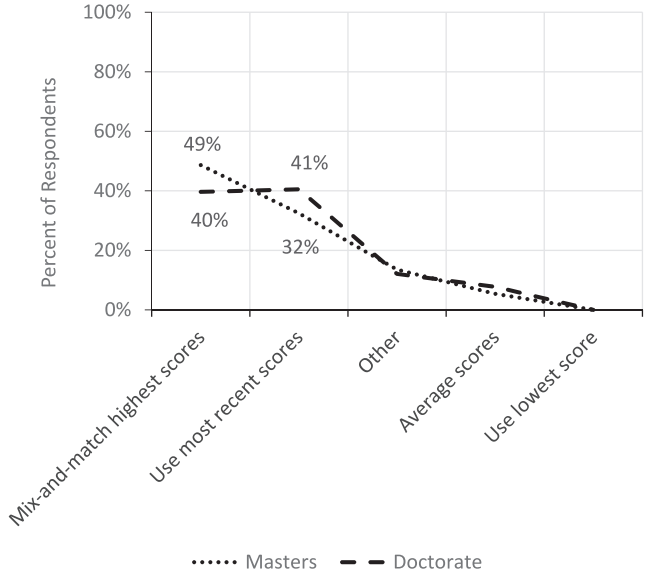

Figure 35 How admissions committees use multiple sets of GRE scores, by degree level.

from each GRE section. Seventeen respondents provided comments about receiving multiple sets of test scores from the same applicant: Four said they would use the highest scores; seven said they would look at all of the scores; two indicated no set policy or procedure; two remarked that if the discrepancy were large enough between different test administrations, they would seek an explanation; and two stated they did not know what they would do or that they had never encountered the issue.

With respect to discipline, STEM participants were less likely than AH participants and SSc participants to use the most recent scores (33% vs. 46% and 47%, RRRs = 0.7) but more likely to use the highest GRE score from each section (48% vs. 32% and 35%, RRRs = 1.5 and 1.3 respectively). See Figure 34.

With respect to degree level, respondents from master's programs were more likely to report mixing and matching the highest scores from different administrations (49% vs. 40%, RRR = 1.2) whereas respondents from doctorate programs were more likely to report that they use the most recent scores (41% vs. 32%, RRR = 1.3). See Figure 35.

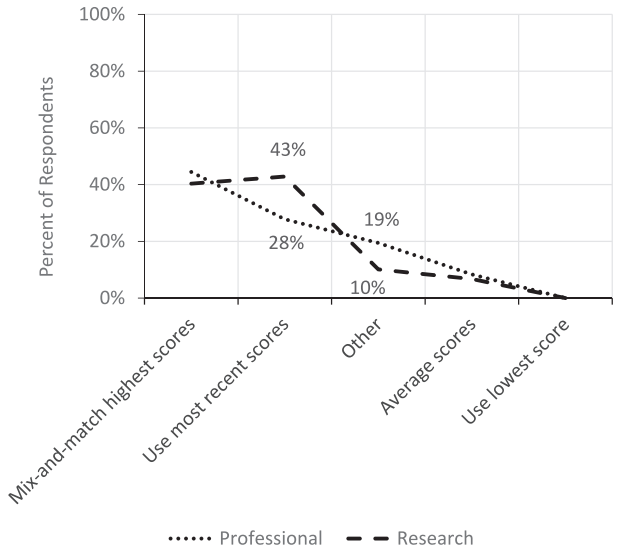

Figure 36 How admissions committees use multiple sets of GRE Scores, by career track.

With respect to differences by career track, participants from professional-track programs were less likely than participants from research-track programs to use the most recent GRE scores submitted to them (28% vs. 43%, RRR = 0.7) and more likely to use other strategies (19% vs. 10%, RRR = 1.9). See Figure 36.

How Are Test Scores From the Old and revised GRE for the Same Applicant Handled?

Figures 37 through 40 illustrate results pertaining to how admissions committees handle test scores for both the old GRE and the revised GRE from a single applicant.

By and large, when faced with GRE test scores from the old version of the GRE for some candidates and from the revised GRE for other candidates, half of respondents said they use score percentile ranks to convert scores. The guidelines regarding the use of GRE scores (ETS, 2015b) discourage this practice, because the percentile ranks for the former and current versions of the GRE General Test are based on different populations. Twenty-four percent of respondents said they do not convert scores at all. Twelve percent indicated that they convert scores from the new GRE to the prior version. That approach is inconsistent with a guideline that encourages conversion from the former to the new scale. By going from the old to the new scale for the verbal and quantitative subtests, one moves from a scale with 61 possible point values (200 – 800, in 10-point increments) to a scale with 41 possible point values (130 – 170, in 1-point increments); more uncertain interpolation in score conversion occurs when going in the opposite direction, from the new scale to the old scale. Twelve percent of respondents revealed that they convert scores from the prior version to the revised GRE, 3% acknowledged that they do something else, and no respondent indicated converting all scores into a brand-new scale score (see Figure 37). Of the four participants who said they do something else when receiving old and revised GRE scores from the same applicant, two stated that they "tend to use percentiles scores" or "no one has any idea what a 'good' raw score is anymore, so we ignore them all in favor of the percentile score." One respondent remarked that the respondent did not know the program's policy for this, and the last participant wrote, "As a program, we don't do anything. Each individual faculty member interprets the scores in whatever way he prefers."

With respect to differences by discipline (Figure 38), STEM and SSc respondents were more likely to say they convert scores from the revised GRE to the prior version, relative to AH respondents (16% and 14% vs. 3%, RRR = 6.1 and 4.7 respectively). As mentioned previously, this directionality in score conversion is inconsistent with GRE guidelines (ETS, 2015b). SSc respondents, relative to AH respondents, were less likely to say they do not convert scores at all (14% and 23% vs. 41%, RRR = 0.3 and 0.6, respectively).

Figure 39 depicts differences by degree level. When faced with converting between old and revised GRE scores, master's program participants were less likely than doctoral program participants to use score percentile ranks and thus be more consistent with GRE guidelines (41% vs. 54%, RRR = 0.8), and master's participants were more likely than doctoral

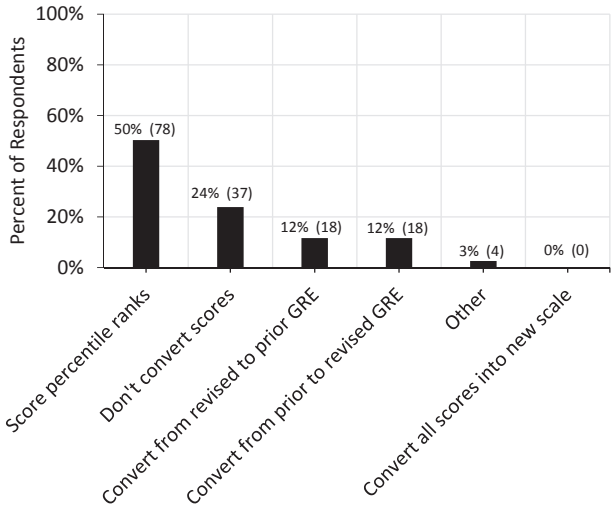

Figure 37 How scores from different versions of the GRE test are handled.

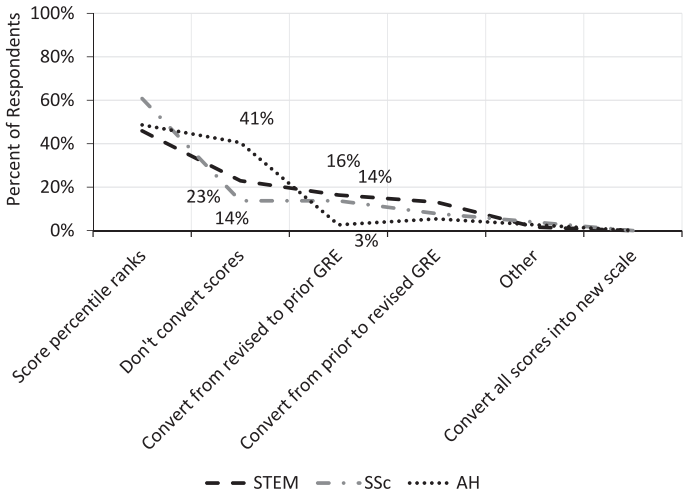

Figure 38 How scores from different versions of the GRE test are handled, by discipline.

participants to not convert scores at all (32% vs. 21%, RRR = 1.6). The groups were equally unlikely to convert between the old and new score scale or to convert all scores to a brand-new scale.

With respect to differences by career track, participants from a professional-track program, relative to those from a research track, were more likely to report that they do not convert scores between prior and recent versions of the GRE test (39% vs. 19%, RRR = 2.0; see Figure 40). Participants from a professional-track program were less likely to report they use score percentile ranks (53% vs. 42%, RRR = 1.3; see Figure 40), and thus they are more likely to be in accord with GRE guidelines concerning score use.

Of those respondents who said they do not convert scores at all when faced with scores from different versions of the test, 81% indicated they use scores from both versions of the test, whereas 19% said they use only scores from the current version of the test (see Figure 41). Of those who said they use scores from both versions of the test ($n = 29$), 89% said they consider different versions of the test separately. (It was unclear to us how that process would work effectively.) Four percent each said they (a) combine or average scores by weighting each the same, (b) combine or average the scores by weighting the current version of the test more, or (c) combine or average the scores by weighting the prior version of the test more (see Figure 42).

With respect to differences by discipline, STEM and SSc programs are less likely than AH programs to use scores from the current version of the test when faced with scores from different versions of the test (14% and 14% vs. 27%, RRR = 0.5

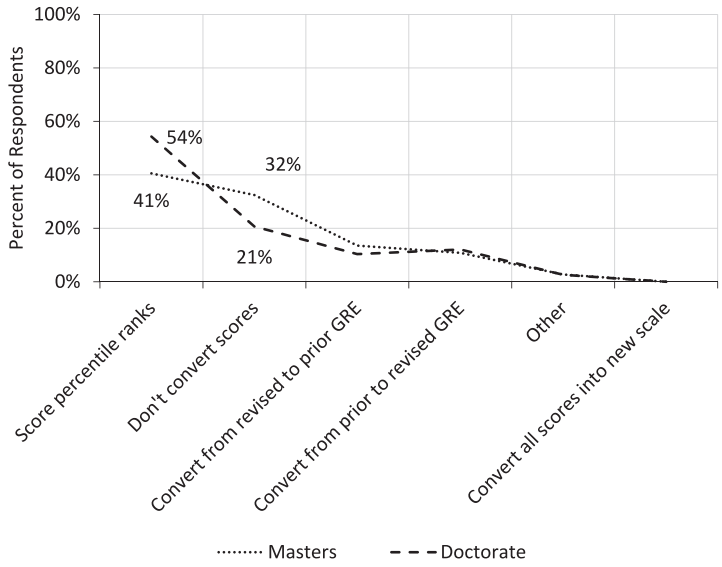

Figure 39 How scores from different versions of the GRE test are handled, by degree level.

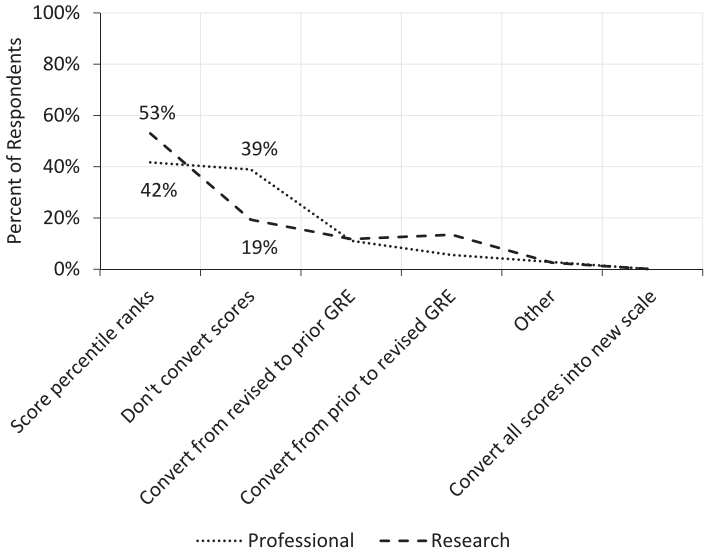

Figure 40 How scores from different versions of the GRE test are handled, by career track.

for both comparisons). See Figure 43. Of those who said they use scores from both versions of the test, STEM and AH programs are more likely than SSc programs to consider the scores separately (100% and 89% vs. 67%), and no STEM programs would combine scores from the different test versions. Only SSc programs indicated that they would combine the scores and weight one of the test versions more (17%); some AH programs also would combine scores, but the tests would be weighed the same (12%). See Figure 44.

With respect to degree level, respondents from master's programs are more likely than respondents from doctorate programs to use both versions of the test (100% vs. 70%, RRR = 1.4) and less likely to use the current version (0% vs. 30%, RRR = 0). See Figure 45. When using both versions, master's program participants are more likely to consider each test score separately than doctorate program participants (100% vs. 80%, RRR = 1.25). See Figure 46.

With respect to career track, respondents from professional-track programs are more likely than respondents from research-track programs to use both versions of the test (92% vs. 74%, RRR = 1.2) and less likely to use the current version (8% vs. 24%, RRR = 0.3). See Figure 47. When using both versions, professional-track program respondents are more likely to consider each test score separately than research-track program respondents (100% vs. 80%, RRR = 1.25. See Figure 48.

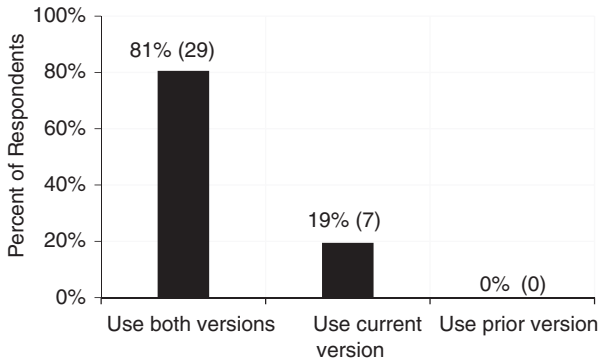

Figure 41 What institutions do if they do not convert scores when faced with different versions of the test.

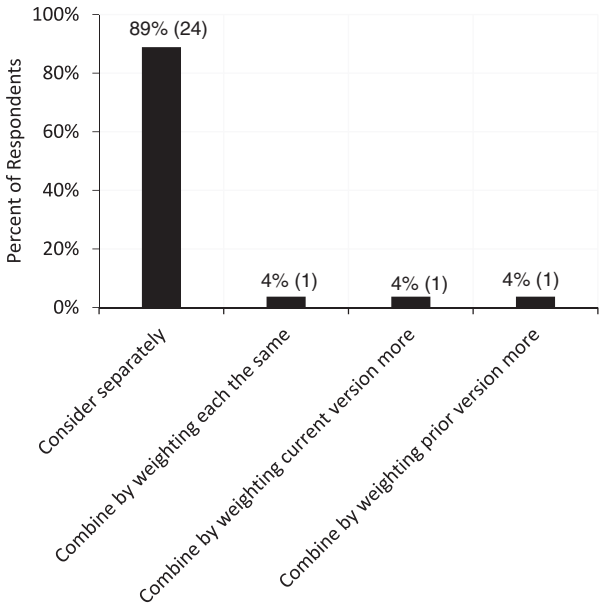

Figure 42 How institutions who use both versions of the GRE handle scores.

Score Conversion Simulations

As mentioned, two of the 11 GRE Program guidelines concerning the use of GRE scores (ETS, 2015b) encourage the use of concordance tables to compare scores from the old and new versions of the GRE General Test. A third guideline cautions against making score comparisons based on percentile ranks from different reference populations. Comparing the percentile ranks for the former GRE test against those for the revised GRE test would be inconsistent with this third guideline. Figures 37 through 40 illustrate that, across different ways of categorizing graduate programs, a plurality or majority of the programs either use percentile ranks or no conversion when faced with GRE scores from different score scales. Another guideline encourages conversion of scores from the former scale to the new scale, but Figures 37 through 40 indicate that some programs convert in the opposite direction. When they forgo conversion of a set of scores, programs tend to somehow consider the differently scaled sets separately; when they favor one scale over the other, it usually is a preference for the newer scale (see Figures 41 – 44). Although the foregoing analyses reveal much about score conversion, we do not know how accurate the decision makers are when they attempt to be fully consistent with the guidelines. That is, when admissions committees attempt to convert scores from the old scale to the new scale using concordance information, how well do they perform? The practicality of the answer to this question diminished beginning in August 2016 (5 years after the launch of the new scale), when ETS will start discontinuing the reporting of scores from the old GRE scale. Nevertheless, the answer provides useful information in the interim as well as for future efforts to help score users convert scores.

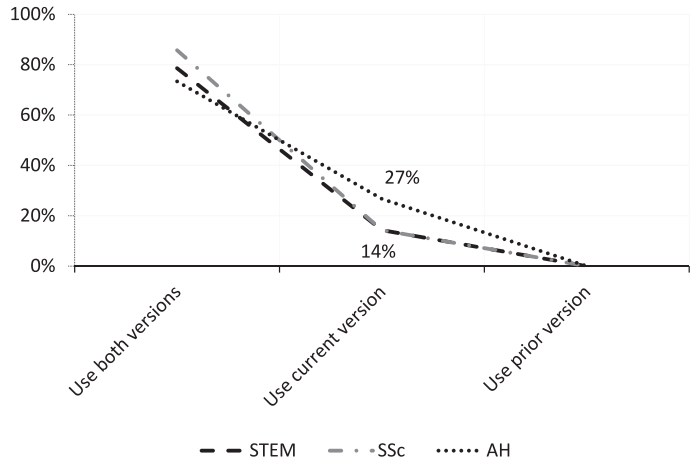

Figure 43 What institutions do if they do not convert scores when faced with different versions of the test, by discipline.

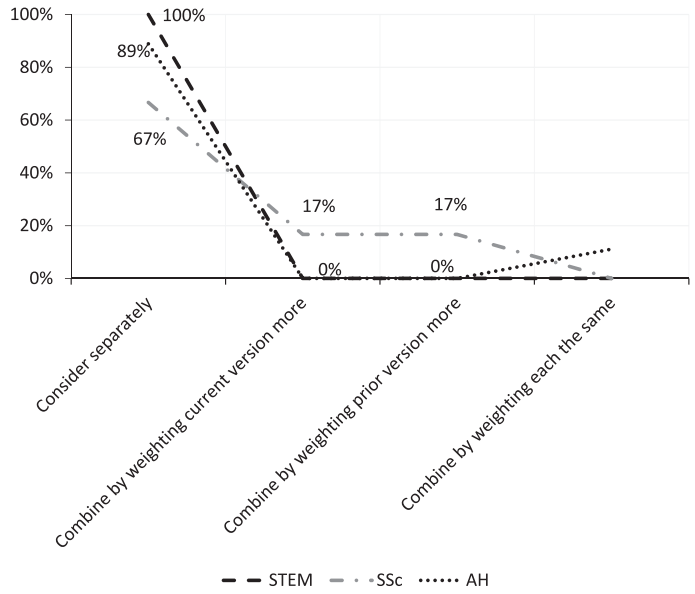

Figure 44 How institutions who use both versions of the GRE handle scores, by discipline.

In order to empirically determine how efficiently admissions committee members compare and utilize that information from concordance tables, participants were given a hypothetical pool of 10 applicants, with five of the applicants having test scores from the old version of the GRE and the other five having test scores from the revised GRE. Question 30 in Appendix A illustrates the task as the respondent would observe it. Along with test scores, participants were given concordance tables that ETS had been providing publicly to score users (see Appendix B) and were asked to rank order the candidates by order of preference for admitting them. Participants were instructed to assume that applicants were equal on all other characteristics relevant for admissions. They were expressly told that ties were possible and that therefore there might be fewer than 10 different rankings. What the participants were unaware of was that the hypothetical pool of applicants had been designed such that each of the five sets of test scores from the old GRE scale was in fact identical to one of the five sets of test scores from the revised GRE scale; it was up to the participants to recognize this fact using the concordance tables and to give tied ranks in their preferences accordingly. In addition, two of the hypothetical candidates (having the same scores but on different scales) had verbal, quantitative, and analytical writing scores that all were higher than the verbal, quantitative, and analytical writing scores for four other hypothetical candidates (two sets of candidates having the same scores but on different scales). Therefore, in addition to the tied ranks previously mentioned, there were eight objectively correct rank orderings. Participants were expected to identify this series of rank order relationships and

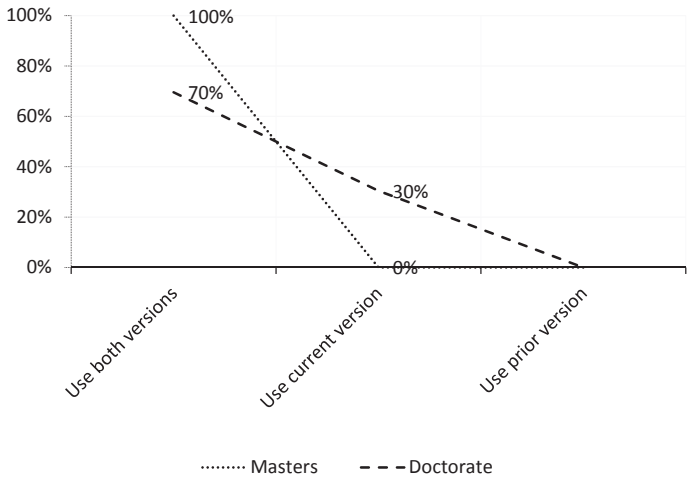

Figure 45 What institutions do if they do not convert scores when faced with different versions of the test, by degree level.

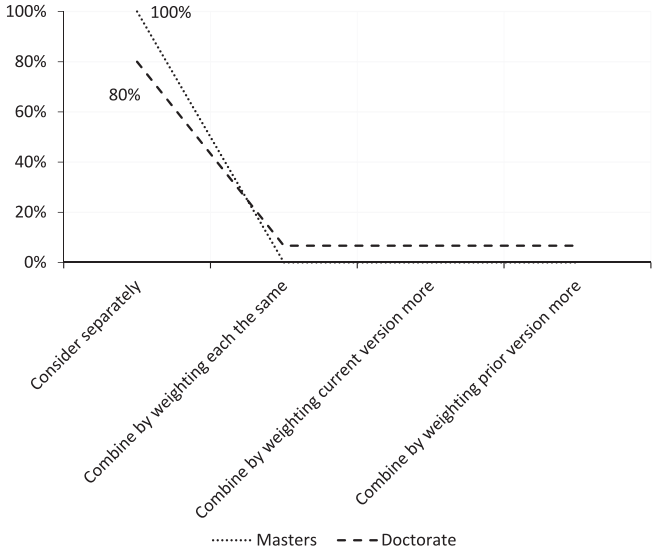

Figure 46 How institutions who use both versions of the GRE handle scores, by degree level.

thus to rank order those two hypothetical candidates as more preferred than the other four. This is the rank ordering referred to below.

Identifying Ties

Overall, few respondents (16%) got all five ties correct, and most (63%) got no ties correct. On average, the mean number of ties correct out of five was 1.3. With respect to differences by discipline, SSc and STEM respondents got more ties correct, relative to AH respondents ($M = 1.7$ and 1.2 respectively vs. 0.6, $d = 0.6$ and 0.4 respectively). With respect to differences by degree level, there was no difference between participants from master's and doctoral programs on the score conversion test ($M = 1.2$ vs. 1.3, $d = 0.0$). With respect to differences by career track, there was no difference between professional and research-track participants on the score conversion task ($M = 1.2$ vs. 1.3, $d = 0.0$).

Rank Ordering

Generally, most respondents (79%) were able to correctly rank order. On average, participants rank ordered correctly in 7.4 out of 8 instances. With respect to differences by discipline, there was no difference by discipline on the rank-order

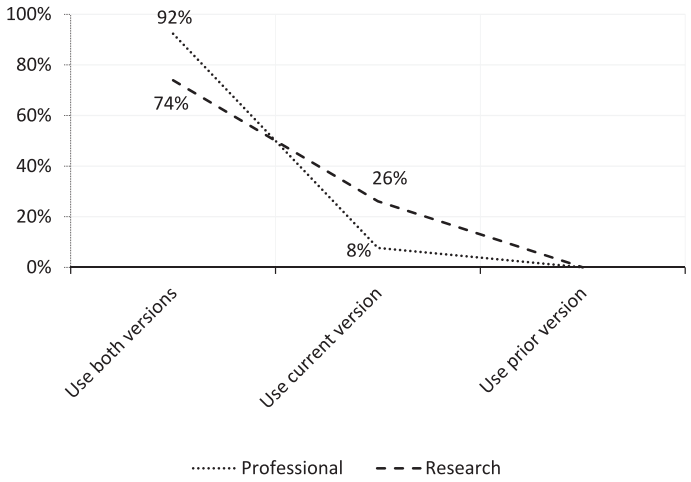

Figure 47 What institutions do if they do not convert scores when faced with different versions of the test, by career track.

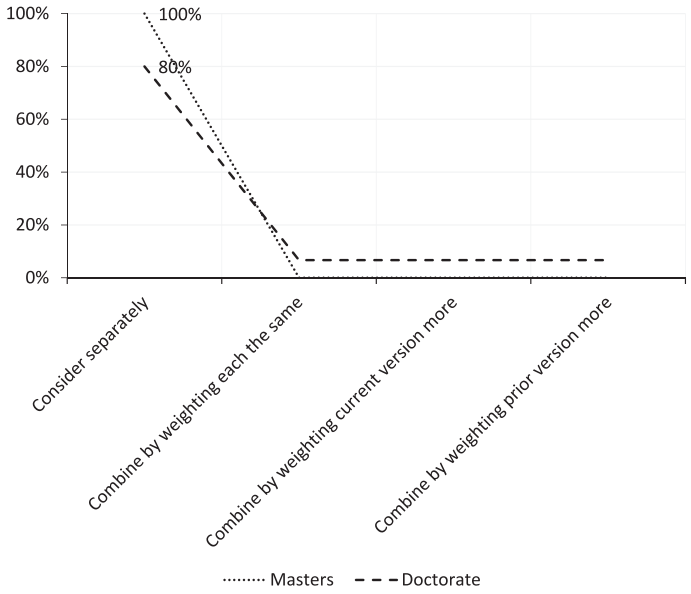

Figure 48 How institutions who use both versions of the GRE handle scores, by career track.

reasoning skills test ($M = 7.4$ vs. 7.5, vs. 7.1; ds ranged from −0.1 to 0.3). With respect to differences by degree level, participants from master's programs performed better than participants from doctoral programs on the rank-ordering reasoning skills test ($M = 7.8$ vs. 7.1, $d = 0.4$). With respect to differences by career track, there was no difference between participants from professional-track and research-track programs on the rank-ordering task ($M = 7.5$ vs. 7.3, $d = 0.1$).

Biases Toward the Old GRE

Because the old GRE scale has been in existence much longer than the revised GRE scale, and because admissions committee members might unconsciously anchor their judgments of scores based on the old scale with larger nominal values (200–800, vs. the new 130–170 range), we wanted to examine if participants who made mistakes with score conversion or rank ordering were biased in favor of old GRE test scores. For the score-conversion test, higher means (closer to 5 on a 0–5 scale) indicate greater average number of biases toward old GRE scores. Lower means (closer to 0 on a 0–5 scale) indicate fewer average number of biases toward old GRE scores. In terms of rank ordering decisions, higher means (closer to 2 on a 0–2 scale) indicate greater average number of biases toward old GRE scores. Lower means (closer to 0 on a 0–2 scale) indicate fewer average number of biases toward old GRE scores.

Eighty-four percent of respondents ($n = 108$) had one or more ties incorrect, and 21% of respondents ($n = 27$) had one or more of the eight rank-orders incorrect. Of those respondents who did not correctly identify tied ranks, STEM respondents showed more of a tendency than SSc respondents did to rank candidates with old GRE scores as better candidates ($M = 3.3$ vs. 2.5, $d = 0.5$). We had expected the opposite, given that most STEM respondents presumably deal with quantitative information on a more regular basis than respondents in other discipline areas. There were no differences between master's and doctoral program participants ($M = 3.0$ vs. 2.9, $d = 0.1$) and professional-track and research-track participants ($M = 3.0$ vs. 2.9, $d = 0.1$) in terms of preference toward the old GRE scale. Of those respondents who did not correctly identify rank ordering, STEM respondents showed more of a tendency than either SSc or AH respondents did in ranking candidates with old GRE scores as better candidates, even though there was no objective reason to do so ($M = 1.33$ vs. .83 and .88, respectively for STEM, SSc, and AH, $d = 0.6$ and 0.6). Again, this was unexpected. There were no differences between master's and doctoral program participants ($M = 1.00$ vs. 1.04, $d = -0.1$) and professional-track and research-track participants ($M = 1.00$ vs. 1.04, $d = 0.0$) in terms of preference for the old GRE scale.

Discussion and Conclusions

Given that the GRE General Test's content, format, and score scale changed as of August 2011, understanding stakeholders' current perceptions and uses of GRE scores is critical to understanding GRE scores' current validity. These changes aside, it is necessary to revisit these perceptions and uses of the GRE General Test because of the passage of time since the prior investigation of this topic (Walpole et al., 2002). Arguably, a similarly comprehensive study has not been conducted since 1984 (Oltman & Hartnett). In the interim, there may have been changes in GRE testing populations; graduate and professional school application, admitted, and enrollee populations; the economy; and the pertinent legal and higher education climates. It is also important to understand the current role of the Subject Tests given their historical integration into the GRE Program and uncertain stakeholder demand for them. This better understanding of current perceptions and uses of the Subject Tests provides valuable context for their validity.

In the context of GRE use, the survey results suggest that general goals of graduate programs do not appear to have changed over the past few decades (see Figures 21 and 22; cf. Oltman & Hartnett, 1984; Powers & Fowles, 2000; Walpole et al., 2002). There are expected differences among program types, such as research programs using the GRE General Test much more than professional programs to award assistantships and fellowships (see Figures 23 – 28). Programs continue to care about maximizing achievement and productivity, enhancing program diversity (albeit not via the GRE tests), and promoting a suitable working environment. In attempting to reach admission and funding objectives, GRE scores (general and subject) are not the principal foundation by which graduate admissions committees tend to make acceptance decisions and instead seem to fall within the middle of the pack of admissions information (see Figures 13 – 16). In terms of use, GRE General Test scores were used by a large percentage of programs represented in the survey responses (see Figures 1, 3, 5, and 7). Scores continue to be used as a further consideration for applicants with weaker credentials or as one part of a holistic selection approach in which committees consider all submitted information to develop an overall impression of a candidate (see Figures 9 – 12; cf. Monahan, 1991; Powers & Fowles, 2000; Rem et al., 1987; Walpole et al., 2002).

At least among the programs represented in the survey, use of GRE General Test scores beyond admissions continues to be relatively common. Even after three decades, about half of programs still use the scores to make decisions regarding fellowship and assistantship awards (see Figures 24, 26, 28; cf. Oltman & Hartnett, 1984). Also, admissions committees still tend to ignore proscriptions of the GRE Board guidelines against the use of cutoffs to filter out applicants (see Figures 9 – 12; cf. Oltman & Hartnett, 1984; Walpole et al., 2002). One unexpected finding was that even though STEM programs still tend to emphasize more strongly quantitative scores relative to SSc and AH programs, there was no appreciable difference across discipline areas for verbal (or analytical writing) scores (see Figure 14; cf. Walpole et al., 2002, which reported STEM programs' lesser emphasis on verbal scores; note that the analytical writing subtest was novel in 2002). A possible explanation is an increased perception about STEM programs that verbal ability is important to success even in STEM programs and fields. Also or instead, STEM programs might see less variability in applicants' quantitative scores, making verbal scores more useful in distinguishing among the applicants. Findings suggest that Subject Tests are at least as infrequently used as they were back in 1984 (see Figures 2, 4, 6, and 8; cf. Oltman & Hartnett, 1984). Furthermore, scores on Subject Tests are deemed relatively unimportant by and large (see Figures 13 through 16). This finding is consistent with more recent qualitative feedback (see Walpole et al., 2002) and quantitative weightings of the perceived lesser importance of GRE Subject Tests (see Powers & Fowles, 2000).

Given that underrepresented minorities tend to perform lower on the GRE General Test (see, e.g., ETS, 2014; Gallagher et al., 2000), it is not surprising that respondents did not perceive the GRE General Test to be useful to enhance diversity, particularly racial and ethnic diversity (see Figures 21–24). Also, across program types, the relative importance of admissions tools changes when racial and ethnic minority candidates are under consideration (as well as international students; see Figures 17–20). This might at least in part reflect an explicit recognition of the "diversity–validity dilemma" previously discussed, or at least its implicit acknowledgement, given its frequent impact on decision-making (Pyburn et al., 2008). Also, the dilemma may partly explain why admissions committees across program types use holistic methods to promote diversity goals (see Figures 9–12; cf. Foderado, 2009). The decisions in Gratz v. Bollinger (2003) and Grutter v. Bollinger (2003), which permit the use of holistic admissions approaches that consider race but not the use of mechanical admission processes that consider race, might also be influencing this tendency to use holistic methods. Whether that will continue given the increasing restrictions on affirmative action (*Fisher v. University of Texas at Austin*, 2013, 2015; *Schuette v. Coalition to Defend Affirmative Action*, 2014) remains to be seen.

As mentioned, GRE Subject Tests are not generally perceived to be as useful as "other" admissions information (see Figures 13–16; cf. Oltman & Hartnett, 1984; Walpole et al., 2002). Although five out of the seven remaining Subject Tests are for STEM fields (mathematics; biochemistry, cell and molecular biology; biology; chemistry; and physics), STEM programs do not view Subject Tests as especially important to admissions decisions (see Figure 14). This perception is unfortunate because large-scale research has demonstrated that the Subject Tests are especially strong predictors of graduate school performance—even more so than subtests in the former version of the General Test (see Kuncel et al., 2001). This might be due to the Subject Tests, which assess discipline-specific knowledge, reliably measuring motivation to deeply study a particular field (Kuncel et al., 2001). As more skills-based assessments, the subtests of the General Test are less likely to detect this kind of enthusiasm. Only a subset of takers of the GRE General Test take the Subject Tests, and they might represent an especially motivated or able population. Also, due to previous selection for college, assessment in college, and variability in college majors and minors, even if all takers of the General Test took the Subject Tests, there might be more measurable variability among examinees in discipline-specific knowledge and motivation to intensely study a particular field than measurable variability in verbal, quantitative, and analytical writing skills.

Other novel issues that this study addresses are related to the multiple sets of GRE General Test scores that graduate and professional programs inevitably receive. Given recent changes to the General Test that became operational in August 2011, admissions committees could receive scores for the former GRE General Test, the revised General Test, or both. When faced with multiple sets of scores (regardless of whether they are on the new or former General Test score scale), admissions committees by and large and across program types either use the highest subtest score from each test administration or just the most recent set of scores (see Figures 33–36). One resulting issue is the extent to which changes in the General Test's score scale—which from the perspective of admissions committees is arguably the most salient change to the General Test and one that admissions committees must address—has changed the way that the GRE General Test is used in admissions. In general and across program types, the change in the scale has not changed the degree to which the General Test is used to make decisions (see Figure 29). Furthermore, in general and across program types, usually there is no greater emphasis placed on scores on the new scale versus the old scale in making admissions or funding decisions; to the very limited extent that there is a preference, it typically is for the new scale (see Figures 30–33). In fact, to the extent that programs in general and across program types converted scores, usually scores on either scale were converted to percentile ranks (presumably by simply looking at the percentile ranks provided on the score reports). See Figures 37–40. To the degree that conversions were made directly from the 200–800 former scale to the 130–170 revised scale or vice versa, there is no preference for one direction versus the other (see Figures 37–40). Given that they have a less quantitative focus, it was unsurprising that arts and humanities programs reported converting scores less often than did social science and STEM programs (see Figure 38).

A limitation to this study is that self-reports of how information is viewed and used are not necessarily how it is actually viewed and used. One can attribute differences to several possible reasons, both conscious and nonconscious. There is empirical evidence that indicates a discrepancy between the weightings ascribed (postdecision) to graduate school admissions criteria and those weightings actually utilized during the decision-making process (Powers & Fowles, 2000). Although we do run a simulation for the use of concordance tables to rank order candidates with scores from both the old and current GRE score scales, the majority of our data are self-reported perceptions. Therefore, although other methods

also would have limitations, we encourage additional investigation of the interpretation and use of GRE scores via other methods (e.g., simulation with policy capturing, third-party observation of actual decision-making, etc.).

Although admissions committees have been forced to adapt to changes to the GRE General Test score scale and the reduction in the number of Subject Tests, this study suggests that perceptions and uses of the General and Subject Tests have not changed a great deal over the last three decades. Admissions, funding, and other decisions affecting graduate and professional students are based on intentional goals of, and nonconscious influences on, programs and institutions. Ultimately, value judgments and other factors determine these goals and influences, and all decisions based on those goals and influences will have tradeoffs. Consequently, we do not evaluate whether the adaptations and changes, or lack thereof, are good or bad, right or wrong. However, we do wish to note that large-scale studies show that cognitive assessments like such as the GRE revised General Test and Subject Tests are robust predictors of performance, both absolutely and relatively (Klieger et al., 2014; Kuncel et al., 2001; Kuncel & Hezlett, 2007; Liu et al., in press). We therefore encourage graduate and professional programs to at least consider these empirical findings when evaluating how to view and use the GRE revised General Test and Subject Tests.

References

Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions*. Princeton, NJ: Princeton University Press.

Briel, J., & Michel, R. (2014). Revisiting the GRE General Test. In C. Wendler & B. Bridgeman (Eds.), with C. Ezzo, *The research foundation for the GRE® revised General Test: A compendium of studies* (p. 1.1.1). Princeton, NJ: Educational Testing Service.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*(2), 75–100.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. Anders Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). Cambridge, MA: Cambridge University Press. Retrieved from http://learnlab.org/uploads/mypslc/publications/chi%20two%20approaches%20chapter%202006.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Conrad, L, Trismen, D., & Miller, R. (Eds.). (1977). *Graduate Record Examinations technical manual.* Princeton, NJ: Educational Testing Service.

Educational Testing Service. (2010). *GRE guide to the use of scores 2010–2011*. Retrieved from https://www.ets.org/s/gre/pdf/2010-11_gre_guide.pdf

Educational Testing Service (2014). *A snapshot of the individuals who took the GRE revised General Test.* Retrieved from https://www.ets.org/s/gre/pdf/snapshot_test_taker_data_2014.pdf

Educational Testing Service. (2015a). *About the GRE® Subject Tests*. Retrieved from https://www.ets.org/gre/subject/about

Educational Testing Service. (2015b). *GRE guide to the use of scores 2015–2016*. Retrieved from http://www.ets.org/s/gre/pdf/gre_guide.pdf

Enright, M. K., & Gitomer, D. (1989). *Toward a description of successful graduate students* (GRE Board Research Report No. 85-17R). Princeton, NJ: Educational Testing Service.

Fisher v. University of Texas at Austin, 570 U.S. __ (2013), U.S. Supreme Court Docket No. 14-981 (2015).

Fix, M., & Struyk, R. J. (Eds.). (1993). *Clear and convincing evidence: Measurement of discrimination in America*. Washington, DC: Urban Institute Press.

Foderado, L. W. (2009, November 1). The whole applicant. *The New York Times.* Retrieved from http://www.nytimes.com/2009/11/01/education/edlife/01admission-t.html

Gallagher, A., Bridgeman, B., & Cahalan, C. (2000). *The effect of computer-based tests on racial/ethnic, gender, and language groups* (GRE Board Report No. 96-21P). Princeton, NJ: Educational Testing Service.

Gratz v. Bollinger, 539 U.S. 244 (2003).

Grutter v. Bollinger, 539 U.S. 306 (2003).

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: Praeger.

Katzenbach, N. D., & Marshall, B. (1998, February 22). Not colorblind: Just blind. *The New York Times*, sec. 6.

Klieger, D. M., Cline, F. A., Holtzman, S. L.,Minsky, J. L., & Lorenz, F. (2014). *New perspectives on the validity of the GRE® General Test for predicting graduate school grades* (GRE Board Research Report No. 14–03). Princeton, NJ: Educational Testing Service. doi:10.1002/ets2.12026

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, *315*, 1080–1081.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162–181.

Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology, 98*(6), 1060 – 1072.

Kuncel, N. R., Klieger, D. M., & Ones, D. S. (2014). In hiring, algorithms beat instinct. *Harvard Business Review, 92*, 32.

Kyllonen, P. C. (2008). *The research behind the ETS Personal Potential Index (*PPI). Princeton, NJ: Educational Testing Service.

Liu, L., Klieger, D., Bochenek, J., Holtzman, S., & Xu, J. (2016). An investigation of the use and predictive validity of GRE scores in a Singaporean university. (GRE Research Report No. GRE-16-01). Princeton, NJ: Educational Testing Service.

Monahan, T. C. (1991) *Using Graduate Record Examination scores in the graduate admissions process at Glassboro State College* (ERIC No. 329183). Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED329183

Oltman, P. K., & Hartnett, R. T. (1984). *The role of GRE General and Subject Test scores in graduate program admission* (GREB Research Report No. 81-8R). Princeton, NJ: Educational Testing Service.

Powers, D. E., & Fowles, M. E. (2000). *Likely impact of the GRE writing assessment on graduate admissions decisions* (GREB Research Report No. GREB-97-06R). Princeton, NJ: Educational Testing Service.

Pyburn, K. M., Ployhart, R. E., & Kravitz, D. A. (2008). The diversity – validity dilemma: Overview and legal context. *Personnel Psychology, 61*(1), 143 – 151.

Rem, R. J., Oren, E. M., & Childrey, G. (1987). Selection of graduate students in clinical psychology: Use of cutoff scores and interviews. *Professional Psychology Research and Practice, 18*(5), 485 – 488.

Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal, 112*, 480 – 518.

Ross, R. P. (1990). Consistency among school psychologists in evaluating discrepancy scores: A preliminary study. *Learning Disability Quarterly, 13*(3), 209 – 219.

Sackett, P. R., and Wilk, S. L. (1994) Within-group norming and other forms of score adjustment in pre-employment testing. *American Psychologist, 49*, 929 – 954.

Schneider, A. M., & Briel, J. B. (1990). *Validity of the GRE: 1988 – 89 summary report* (GREB Research Report No. GREB-90-01VSS). Princeton, NJ: Educational Testing Service.

Schuette v. Coalition to Defend Affirmative Action, 572 U.S. __ (2014).

Walpole, M., Burton, N. W., Kanyi, K., & Jackenthal, A. (2002). *Selecting successful graduate students: In-depth interviews with GRE users* (GREB Research Report No. GREB-99-11R). Princeton, NJ: Educational Testing Service.

Appendix A

Survey

1. Please provide the following information
Last name: _____
First name: _____
Middle name (optional): _____
Current Title: _____
Current Institution: _____
Current Department: _____
Current Program: _____

2. Have you served (or do you serve) on an admissions committee? (Choose all that apply)
□ Served as a chair of the committee → GO TO 3.
□ Served as a committee member → GO TO 4.
□ I did not serve on an admissions committee, but provided advice to the committee regarding applicants → GO TO 5.
□ I did not serve on an admissions committee, but the committee provided advice to me regarding applicants → GO TO 5.
□ I have not served on an admissions committee, nor have I received or given advice to such a committee → EXIT SURVEY.

3. For how many years did you serve as a chair of the admissions committee?
# of years: _____

4. For how many years did you serve as a member of the admissions committee?
# of years: _____

5. If different from your current institution, department and program, please describe the institution, department, and program (e.g., Harvard University Clinical Psychology) in which you have been most involved making admissions decisions. **State only one institution, department, and program even if you were involved in more than one:**
_____
_____

6. For which type of graduate degree program have you been most responsible for making admissions decisions? **Choose only one even if you have been or are responsible for more than one:**
□ Professional master's
□ Research master's
□ Professional doctorate
□ Research doctorate
□ Other (Please specify): _____

PLEASE ANSWER ALL OF THE QUESTIONS THAT FOLLOW SPECIFICALLY FOR THE GRADUATE PROGRAM THAT YOU CHOSE IN THE PREVIOUS QUESTION.

SHOW IF (Served as a committee member) = Selected

7. Which statement best describes your admissions policy?

□ Essentially "open door" (i.e. non-selective)
□ Somewhat competitive (i.e. moderate credentials acceptable; may not always consider undergraduate grades or test scores)
□ Moderately competitive (i.e. strong credentials required)
□ Very competitive (i.e. some strong candidates rejected)
□ Extremely competitive (i.e. only exceptional candidates accepted)

8. Rate the criteria below in terms of their importance for admissions decisions.

Importance

	Not used	Not very important	Moderately important	Very important	Extremely important
GRE Verbal Reasoning score	□	□	□	□	□
GRE Quantitative Reasoning score	□	□	□	□	□
GRE Analytical Writing score	□	□	□	□	□
GRE Subject Test (e.g., Chemistry, Psychology, etc.) score most appropriate to **your academic unit**	□	□	□	□	□
GRE Subject Test score most appropriate to the applicant's undergraduate major	□	□	□	□	□
Particular **subscores** on the appropriate GRE Subject Test	□	□	□	□	□

Comments (if any):_____
_____

9. Rate the criteria below in terms of their importance for admissions decisions.

Importance

	Not used	Not very important	Moderately important	Very important	Extremely important
Quality of undergraduate institution	☐	☐	☐	☐	☐
Undergraduate grade point average in junior and senior years	☐	☐	☐	☐	☐
Undergraduate grade point average in major field	☐	☐	☐	☐	☐
Undergraduate grade point average overall	☐	☐	☐	☐	☐
Undergraduate major related to field of graduate study	☐	☐	☐	☐	☐
Recommendation letters from faculty **not known** by the admissions committee	☐	☐	☐	☐	☐
Recommendation letters from faculty **known** by the admissions committee	☐	☐	☐	☐	☐
Non-faculty recommendations	☐	☐	☐	☐	☐
Scores from the *ETS*® Personal Potential Index (PPI)	☐	☐	☐	☐	☐
Scores from another assessment of non-cognitive skills (i.e. other than PPI).	☐	☐	☐	☐	☐
Test of English proficiency for international applicants.	☐	☐	☐	☐	☐
Other test scores.	☐	☐	☐	☐	☐
Familiarity (i.e. applicant known by your faculty)	☐	☐	☐	☐	☐
Other academic achievements (presentations, publications, projects)	☐	☐	☐	☐	☐
Personal interview	☐	☐	☐	☐	☐
Personal statement	☐	☐	☐	☐	☐
Educational or career aspirations of applicant	☐	☐	☐	☐	☐
Prior research experience	☐	☐	☐	☐	☐
Prior laboratory experience	☐	☐	☐	☐	☐
Prior work experience (while attending school or after graduation)	☐	☐	☐	☐	☐
Having spent time gaining life experience or practical experience after graduating from college	☐	☐	☐	☐	☐
Other criteria (Please list/describe in Comments)	☐	☐	☐	☐	☐

Comments (if any):_____
_____

10. Check the box beside any group below if the relative importance of the criteria listed above is different for that group.

- ☐ International applicants
- ☐ Members of under-represented racial/ethnic groups
- ☐ Older applicants
- ☐ Applicants with a post-baccalaureate degree
- ☐ Female applicants
- ☐ Applicants with disabilities
- ☐ Other group (Please specify): _____

11. If you selected any of the options listed above, please explain how the importance of the criteria differs:

_____
_____

IF
GRE Verbal Reasoning score = Not used AND
GRE Quantitative Reasoning score = Not used AND
GRE Analytical Writing score = Not used AND
GRE Subject Test (e.g., Chemistry, Psychology, etc.) score most appropriate to your academic unit] = Not used AND
GRE Subject Test score most appropriate to the applicant's undergraduate major = Not used AND
Particular subscores on the appropriate GRE Subject Test = Not used
THEN → GO TO 31.

IF
GRE Subject Test (e.g., Chemistry, Psychology, etc.) score most appropriate to your academic unit] = Not used AND
GRE Subject Test score most appropriate to the applicant's undergraduate major = Not used AND
Particular subscores on the appropriate GRE Subject Test = Not used
THEN → GO TO 16.

12. Do you have a formal policy regarding admissions for students with disabilities?

- ☐ Yes
- ☐ No

13. If yes, please explain and/or copy and paste the policy statement:

_____
_____

IF
GRE Verbal Reasoning score = Not used AND
GRE Quantitative Reasoning score = Not used AND
GRE Analytical Writing score = Not used AND
THEN → GO TO 14.

14. What is your policy regarding GRE **Subject** Test scores?
- ☐ Required for all applicants
- ☐ Required but can be waived in some circumstances.
- ☐ Recommended for all applicants
- ☐ Required or recommended for **some** degree applicants
- ☐ Not required or recommended but will consider if submitted

15. Please feel free to specify the circumstances and applicants for which you require or recommend GRE **Subject** Test scores:

_____
_____

IF
GRE Verbal Reasoning score = Not used AND
GRE Quantitative Reasoning score = Not used AND
GRE Analytical Writing score = Not used AND
THEN → GO TO 31.

16. What is your policy regarding GRE **General** Test (Verbal Reasoning, Quantitative Reasoning, and Analytical Writing) scores?
- ☐ Required for all applicants
- ☐ Required but can be waived in some circumstances.
- ☐ Recommended for all applicants
- ☐ Required or recommended for **some** degree applicants
- ☐ Not required or recommended but will consider if submitted

17. Please feel free to specify the circumstances and applicants for which you require or recommend GRE **General** Test scores:

_____
_____

18. How important is the role of the GRE **General** Test in achieving the following goals?:

Importance

	Not applicable	Not very important	Moderately important	Very important	Extremely important
Matching applicants' research interests with faculty expertise	☐	☐	☐	☐	☐
Minimizing time-to-degree completion	☐	☐	☐	☐	☐
Selecting applicants who will be able to handle graduate school coursework.	☐	☐	☐	☐	☐
Selecting students who will learn important skills outside of the classroom (e.g., ability to run special software or lab equipment that facilitates research).	☐	☐	☐	☐	☐
Selecting applicants who will be good teaching assistants or instructors	☐	☐	☐	☐	☐
Selecting applicants whose personality and interests seem to be a "good fit" with the culture of the field	☐	☐	☐	☐	☐
Selecting applicants who will promote a cohesive environment in the program/department/laboratory	☐	☐	☐	☐	☐
Maximizing gender diversity	☐	☐	☐	☐	☐
Maximizing racial/ethnic diversity	☐	☐	☐	☐	☐
Maximizing the diversity of the admitted class of students based on characteristics other than race/ethnicity or gender	☐	☐	☐	☐	☐
Satisfying financial needs of the program or institution	☐	☐	☐	☐	☐
Meeting other goals - (Please describe in Comments)	☐	☐	☐	☐	☐

Comments (if any):_____
_____

19. How do you use GRE scores in admissions? (Choose all that apply)
☐ Applicants scoring below a specific score are not considered.
☐ GRE scores are used to categorize applicants (e.g., "probable," "possible," "unlikely") before other credentials of an applicant are reviewed.
☐ When the applicant's other credentials are strong, test scores are unimportant. For applicants with weaker credentials, the test scores are expected to compensate.
☐ Points are assigned to each applicant's test scores and other credentials based on how important they are believed to be. The points are summed, and the applicants with the highest sums are offered admission.

□ Prediction formulas (e.g., regression equations), based on test scores and other credentials of previous applicant groups, are used for selecting among new applicants for admission.
□ GRE scores are used as one factor in a holistic review of applicants' files. For example, an admissions committee might look at undergraduate GPA, GRE scores, and personal statements and make a decision whether to admit a candidate.
□ Other (Please specify): _____

20. Do you use GRE scores for (select all that apply):

	GRE General Scores	GRE Subject Scores
Academic advising?	□	□
Awarding assistantships and fellowships?	□	□
Placement of students in courses?	□	□
Comprehensive examination or other graduation requirement?	□	□
Another purpose? (Please specify):	□	□

21. If you selected Another purpose, please specify:

_____
_____

22. Are you familiar with the guidelines that the GRE Board has established for the use of GRE scores?
□ Yes
□ No

23. How do you use multiple sets of GRE scores that an applicant sends?
□ Use the most recent scores
□ Use the highest score from each GRE section across test administrations (i.e. "mix-and-match" highest scores);
□ Use the lowest score from each GRE section across test administrations (i.e. "mix-and-match" lowest scores);
□ Average the scores from each GRE section across the test administrations
□ Other (Please specify): _____

24. Scores on the Verbal Reasoning and Quantitative Reasoning sections of the GRE revised General Test (administered after August 1, 2011) range from 130 to 170, in 1-point increments. Scores on the prior version of the test range from 200 to 800, in 10-point increments.

How has the new scale affected your use the of GRE scores?
□ It has not changed the way we use GRE scores
□ GRE scores play a bigger role in admissions
□ GRE scores play a smaller role in admissions
□ Other (Please specify): _____

25. Feel free to provide any additional comments regarding changes to the score scale:

_____

_____

26. It is possible that a single applicant will submit GRE scores on the current version of the GRE test (scored 130 to 170) as well as scores on the prior version of the test (scored 200 to 800). It also is possible that some applicants will submit GRE scores on the current version of the test (scored 130 to 170) and that other applicants will submit scores on the prior version of the test (scored 200 to 800). In either event, you:

- Always give preference to scores from the new version of the test (scored 130 to 170);
- Sometimes give preference to scores from the new version of the test (scored 130 to 170);
- Do not let the version of the test affect admissions or funding decisions;
- Sometimes give preference to scores from the prior version of the test (scored 200 to 800);
- Always give preference to scores from the prior version of the test (scored 200 to 800).

27. Do you convert GRE scores on the current version of the test (scored 130 to 170) into scores on the prior version of the test (scored 200 to 800), vice-versa, neither, or do you do something else when faced with scores from different versions of the test? (Please note that ETS has published concordance tables to help decision makers to convert scores from one scale to the other.)

- We convert GRE scores on the current version of the test (scored 130 to 170) into scores on the prior version of the test (scored 200 to 800). → GO TO 30.
- We convert GRE scores on the prior version of the test (scored 200 to 800) into scores on the current version of the test (scored 130 to 170). → GO TO 30.
- We convert all scores into a brand new score scale (neither a 130 to 170 nor 200 to 800 score scale). → GO TO 30.
- We just use score percentile ranks (which can be thought of as a type of score conversion). GO TO 30.
- We do not convert scores at all. → GO TO 28.
- Other (Please specify): _____ →GO TO 30.

28. If you do not convert scores at all, what do you do?

- We only use scores from the current version of the test → GO TO 30.
- We only use scores from the prior version of the test → GO TO 30.
- We use scores from both versions of the test → GO TO 29.

29. If you use scores from both versions of the test, how do you treat the scores?

- Consider scores on different versions of the test separately
- Combine or average them by weighting each the same
- Combine or average them by weighting the scores on the current version of the test more
- Combine or average them by weighting the scores on the prior version of the test more

30. Below is a hypothetical pool of applicants, along with their GRE scores. Applicants have submitted scores from both the prior and current versions of the GRE test. Rank order the candidates by order of preference, with 1 meaning most preferred and 10 meaning least preferred. Ties are permitted. (If there are any ties, then the lowest-ranked candidate below would be ranked higher than 10.) Assume that applicants are equal on all other relevant characteristics. The score scale for the Analytical Writing section has not changed. The Analytical Writing section continues to be scored 0 to 6, in 0.5 point increments. To see the concordance tables showing what a score on one scale is equivalent to on the other scale, click here

Rank Order 1 to 10 (1 = most preferred; 10 = least preferred)

Applicant 1: Verbal Reasoning=169, Quantitative Reasoning=161, Analytical Writing=4	
Applicant 2: Verbal Reasoning=240, Quantitative Reasoning=740, Analytical Writing=2.5	
Applicant 3: Verbal Reasoning=560, Quantitative Reasoning=650, Analytical Writing=3	
Applicant 4: Verbal Reasoning=740, Quantitative Reasoning=770, Analytical Writing=4	
Applicant 5: Verbal Reasoning=800, Quantitative Reasoning=540, Analytical Writing=5	
Applicant 6: Verbal Reasoning=134, Quantitative Reasoning=147, Analytical Writing=5.5	
Applicant 7: Verbal Reasoning=170, Quantitative Reasoning=145, Analytical Writing=5	
Applicant 8: Verbal Reasoning=131, Quantitative Reasoning=158, Analytical Writing=2.5	
Applicant 9: Verbal Reasoning=270, Quantitative Reasoning=580, Analytical Writing=5.5	
Applicant 10: Verbal Reasoning=157, Quantitative Reasoning=151, Analytical Writing=3	

Comments (if any):_____

_____

GO TO 32.

31. How important is each of the listed reasons for your not requiring or recommending GRE test scores for admissions? (These are based on a belief or perspective that an institution, department, program, committee, or individual could have.)

	Unimportant	Moderately important	Very important
Test scores do not predict performance well	☐	☐	☐
The test poses special difficulties for those with disabilities	☐	☐	☐
Test scores do not predict equally well across demographic groups	☐	☐	☐
Test scores are redundant with other information we receive	☐	☐	☐
Other information received from the applicant is more important	☐	☐	☐
Almost all applicants are admitted	☐	☐	☐
The GRE tests seem to penalize the better, more creative students in our academic unit	☐	☐	☐
GRE Subject Test content is not appropriate for our use	☐	☐	☐
Requiring or recommending GRE scores would add more work to an already complicated process	☐	☐	☐
Requiring or recommending GRE scores would add more cost to the application process	☐	☐	☐
Other reason. Please specify in comments.	☐	☐	☐

Comments (if any):_____
_____

Background Questions (Optional)

32. Gender:
☐ Female
☐ Male
☐ I prefer not to answer

33. Race/Ethnicity. Please select all that apply:
☐ American Indian or Alaska Native
☐ Asian
☐ Black or African American
☐ Native Hawaiian or Other Pacific Islander
☐ White or Caucasian
☐ I do not identify with any of these options. If you would like to report a different response than one listed above, please do so here: _____
☐ I prefer not to answer

34. Are you of Latino/Hispanic background?

□ Yes
□ No
□ I prefer not to answer

35. What is your age (in years)? If you prefer not to answer this question, you may leave the box empty and proceed to the next question ___

36. Please provide your country or countries of citizenship. You can indicate up to two countries from the drop-down menus below.

□ Abkhazia – Republic of Abkhazia
□ Afghanistan – Islamic Republic of Afghanistan…
……..
□ ……Zambia – Republic of Zambia
□ Zimbabwe – Republic of Zimbabwe
□ I prefer not to answer

37.
□ Abkhazia – Republic of Abkhazia
□ Afghanistan – Islamic Republic of Afghanistan…
……
□ ……Zambia – Republic of Zambia
□ Zimbabwe – Republic of Zimbabwe
□ I prefer not to answer

This Question is Conditionally Hidden if: (35 = United States – United States of America OR36 = United States – United States of America)

38. If you are not a U.S. citizen, are you a U.S. permanent resident?
□ Yes
□ No
□ I prefer not to answer

39. Do you possess significant work experience outside of academia such as in industry, government, or the non-profit sector?
□ Yes
□ No
□ I prefer not to answer

40. Would you be willing to participate in a brief, confidential interview concerning your admissions practice? If so, please enter an e-mail address (for an account that is not shared) and/or telephone number where we may contact you.

E-mail: _____
Telephone number: _____

41. **In order to process your $50 gift card, we need the following information.**

Please select the option that best describes you:
 I am a U.S. citizen/permanent resident. I **have not** received payment from ETS for any other service performed during the current calendar year 2015.
 I am a U.S. citizen/permanent resident. I **have** received payment from ETS during the current calendar year 2015 that will NOT meet or exceed $600 in total and I am not an employee of ETS.
 I am a U.S. citizen/permanent resident. I **have** received payment from ETS during the current calendar year 2015 that WILL meet or exceed $600 in total and I am not an employee of ETS.
 I am a foreign national on a non-immigrant visa. I certify that I **have not** received payments exceeding $100 in total, including this study, from ETS in the calendar year 2015.

42. Upon completion of the survey, please send the gift card to:
First name: _____
Last name: _____
Phone number: _____
Email address: _____

Appendix B

Concordance Tables for Prior and Revised Versions of GRE General Test

ETS GRE

Listening. Learning. Leading.®

Table 1D: Verbal Reasoning Concordance Table

Prior Scale	New Scale	% Rank*
800	170	99
790	170	99
780	170	99
770	170	99
760	170	99
750	169	99
740	169	99
730	168	98
720	168	98
710	167	97
700	166	96
690	165	95
680	165	95
670	164	93
660	164	93
650	163	91
640	162	89
630	162	89
620	161	87
610	160	84
600	160	84
590	159	81
580	158	78
570	158	78
560	157	73
550	156	70
540	156	70
530	155	66
520	154	62
510	154	62
500	153	58

Verbal Reasoning Concordance Table (continued)

Prior Scale	New Scale	% Rank
490	152	53
480	152	53
470	151	49
460	151	49
450	150	44
440	149	40
430	149	40
420	148	36
410	147	32
400	146	28
390	146	28
380	145	24
370	144	21
360	143	18
350	143	18
340	142	15
330	141	12
320	140	10
310	139	7
300	138	6
290	137	5
280	135	2
270	134	2
260	133	1
250	132	1
240	131	1
230	130	
220	130	
210	130	
200	130	

*Based on the performance of all examinees who tested between August 1, 2011 and April 30, 2013. Percentile ranks are updated yearly.

Table 1E: Quantitative Reasoning Concordance Table

Prior Scale	New Scale	% Rank*
800	166	93
790	164	89
780	163	87
770	161	81
760	160	78
750	159	75
740	158	72
730	157	69
720	156	65
710	155	61
700	155	61
690	154	57
680	153	53
670	152	49
660	152	49
650	151	45
640	151	45
630	150	41
620	149	37
610	149	37
600	148	33
590	148	33
580	147	29
570	147	29
560	146	25
550	146	25
540	145	22
530	145	22
520	144	18
510	144	18
500	144	18

Quantitative Reasoning Concordance Table (continued)

Prior Scale	New Scale	% Rank
490	143	15
480	143	15
470	142	13
460	142	13
450	141	11
440	141	11
430	141	11
420	140	8
410	140	8
400	140	8
390	139	6
380	139	6
370	138	5
360	138	5
350	138	5
340	137	3
330	137	3
320	136	2
310	136	2
300	136	2
290	135	2
280	135	2
270	134	1
260	134	1
250	133	1
240	133	1
230	132	
220	132	
210	131	
200	131	

Note: Score users should use special care in evaluating test takers who received a Quantitative Reasoning score at the top end of the prior 200-800 score scale. Now, with the new 130-170 score scale, we can provide more differentiation for higher ability test takers. However, test takers who took the prior test and received an 800 on the Quantitative Reasoning measure, received the highest score possible that they were able to earn on the measure. Therefore, this information should be considered when making admissions decisions.

*Based on the performance of all examinees who tested between August 1, 2011, and April 30, 2013. Percentile ranks are updated yearly.

Suggested citation:

Klieger, D. M., Belur, V., & Kotloff, L. J. (2017). *Perceptions and uses of GRE® scores after the launch of the GRE® revised General Test in August 2011* (GRE Board Research Report No. GRE-17-01). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/ets2.12130

Action Editor: Brent Bridgeman

Reviewers: This report was reviewed by the GRE Technical Advisory Committee and the Research Committee and Diversity, Equity and Inclusion Committee of the GRE Board